

Dissertation

Mathematiklehrkräfte im Implementations- prozess der nationalen Bildungsstandards

zur Erlangung des akademischen Grades

doctor philosophiae (Dr. phil.)

im Fach Erziehungswissenschaften

eingereicht am 28.04.2017

an der Kultur-, Sozial- und Bildungswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von Jenny Frenzel

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekanin der Kultur-, Sozial- und Bildungswissenschaftlichen Fakultät

Prof. Dr. Julia von Blumenthal

Gutachter/in:

1. Prof. Dr. Hans Anand Pant

2. Prof. Dr. Olaf Köller

3. Prof. Dr. Barbara Asbrand

Datum der Verteidigung: 19.01.2018

Zusammenfassung

Diese Arbeit untersucht das Verhalten und die Einstellungen von Lehrkräften und Schülerinnen und Schüler im beginnenden Implementationsprozess der nationalen Bildungsstandards. Diese top-down eingeführte Bildungsinnovation intendiert eine Umstellung des Unterrichts, der auf die in den Standards formulierten Kompetenzen fokussiert. Zunächst wird herausgearbeitet, was unter einem kompetenzorientierten Mathematikunterricht zu verstehen ist und im Folgenden, wie eine solche Intention in die Praxis überführt werden kann. Dazu werden verschiedene Implementationstheorien und –strategien vorgestellt.

Die für die deutschen Schulen unbeantwortete Frage nach der Vorgehensweise und Umsetzung der Bildungsstandards versuchte das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) durch ein Implementationskonzept zu lösen. Anhand von zwei Studien wurden jeweils zehn Merkmale eines kompetenzorientierten Unterrichts an Lehrkräften und Schülerinnen und Schüler im Längsschnitt analysiert und einzeln ausgewertet. Zu jeder Datenerhebung arbeiteten die Schülerinnen und Schüler zudem an einem kompetenzorientierten Leistungstest. Die Stichproben unterteilten sich in Projekt- und Vergleichsschulen. Die Projektschulen erhielten über mehrere Jahre eine externe Unterstützung in Form einer Setkoordinatorin, die in den Fachkonferenzen einen Input und beratende Hilfestellung bei der Entwicklung und Erprobung von kompetenzorientierten Lehr-Lernarrangements anbot. Angelehnt an das SINUS-Programm wurden weitere zentrale Fortbildungsveranstaltungen angeboten, die den Lehrkräften in den Projektschulen die Möglichkeit bot, sich miteinander auszutauschen.

In einer zweiten Studie konnte das IQB-Implementationskonzept mit dem SINUS-Folgeprojekt „Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften (KOU)“ evaluiert werden.

Im Ergebnis zeigen die Drop-Out-Analysen zunächst, dass Lehrkräfte, die den Bildungsstandards gegenüber positiv eingestellt sind und bereits in einem höheren Grad miteinander kooperieren, Engagement im Projekt zeigen.

Der Unterricht in den Projektschulen erfuhr eine stärkere Orientierung hin zum mathematischen Argumentieren, Modellieren und Probleme lösen.

In den Projektschulen der Studie 1 nehmen die Schülerinnen und Schüler eine verstärkte Kompetenzorientierung im Bereich der Verwendung von mathematischen Darstellungen wahr. Ebenso konnte das IQB-Implementationskonzept positiv evaluiert werden.

Summary of the Doctoral Thesis

This doctoral thesis focuses on teachers' development during the implementation process of the national educational standards. Consider the standards, which were introduced by resolution, classroom activities are focused to improve student learning. To reach the standards, a change in teaching culture is necessary or at least an addition of teachers' classroom activities. Therefore, the concept of competence-orientation especially for mathematics is described and discussed. Further implementation theories and implementation strategies are suggested and discussed.

When educational standards became active, implementation concepts were missing so far. So the Institute for Educational Quality Improvement (IQB) developed an implementation concept and got support by nine schools to test the efficacy. The professional learning communities in mathematics of the schools got external support by a supervisor and input by researchers during their intended change process.

In study 1, ten factors were analysed to investigate teachers' engagements and concerns as well as their classroom activities in a longitudinal design for three times. Students of ninth grade were tested in mathematics and asked about their classroom activities for three times, too. Additionally, eight schools were tested and asked as a control group.

Study 2 added another ten schools with a slightly different implementation concept. So, the IQB-concept could be evaluated. Also in study 2, the ten factors are analysed and differentiated by concept.

Even the drop-out analyses show, teachers with a positive attitude towards the standards and a higher degree in cooperation, support the project and take part in the longitudinal study. Furthermore, their lessons change in aspects of argument, modeling and solving problems mathematically. Students of study 1 observe more activities in working with figures and diagrams during their mathematics classes.

In study 2, the IQB implementation concept could be evaluated successfully by comparing teachers' changes with the development in the other schools.

1	Einleitung	9
1.1	Problemaufriss	11
1.2	Eingrenzung des Untersuchungsgegenstandes	15
2	Kompetenzerwerb auf der Grundlage von Bildungsstandards	18
2.1	Kompetenzen als Ergebnis schulischen Lernens	19
2.1.1	Abgrenzung zum Konstrukt der Intelligenz	21
2.1.2	Performanz und Kompetenz	25
2.1.3	Kompetenzerwerb und Kompetenzaufbau	28
2.2	Die Bildungsstandards für das Fach Mathematik in der Sekundarstufe I	35
2.2.1	Das Kompetenzmodell für das Fach Mathematik	35
2.2.2	Das Kompetenzstufenmodell für das Fach Mathematik	43
2.3	Standard-basierte Unterrichtsgestaltung durch Kompetenzorientierung	50
2.3.1	Unterrichtsverständnis und Unterrichtsskript	52
2.3.2	Kompetenzorientierung im Mathematikunterricht	60
2.3.2.1	Fachlich gehaltvolle Unterrichtsgestaltung	60
2.3.2.2	Kognitive Aktivierung im Mathematikunterricht	65
2.3.2.3	Effektive und schülerorientierte Unterrichtsführung	70
2.3.3	Neue Aufgabenkultur	73
2.4	Zusammenfassung und zentrale Aussagen	80
3	Die Implementation eines standard-basierten Unterrichts	82
3.1	Innovationstheorien im Anwendungsbereich der Bildungsforschung	84
3.1.1	Fullan's „The new meaning of educational change“	84
3.1.2	Sense-making der Agenten	89
3.1.3	Concerns-based-Adoption-Model	92
3.2	Lehrerkognitionen und Potentiale der Veränderung der Unterrichtspraxis in kooperativen Strukturen	97
3.2.1	Unterrichtliches Handeln auf der Grundlage von Einstellungen, Überzeugungen und subjektive Theorien	98
3.2.2	Professionalisierung im Kontext von Unterrichtsentwicklung	105
3.2.3	Professionelle Lerngemeinschaften	111
3.3	Förderliche Faktoren im Prozess der Implementation einer Bildungsinnovation	116
3.3.1	Die Genese von Faktoren aus theoretischer Sicht	116
3.3.2	Nachhaltige Implementation erfolgreicher Projekte	120
3.3.3	Studien aus der Begleitforschung zur Implementation der Bildungsstandards	126
4	Fragestellung	130
5	Die Begleitforschung zur Implementation der Bildungsstandards	134

5.1 Das Projekt zur „Erprobung Konzeptes zur Implementation der Bildungsstandards im Fach Mathematik in Berliner und Brandenburger Schulen“ (Teilstudie 1)	136
5.1.1 Projektbeschreibung.....	136
5.1.2 Modulbeschreibung.....	138
5.1.3 Die Umsetzung der Konzeption	141
5.1.4 Der Ablauf der Begleitforschung	142
5.1.5 Stichprobe der Studie 1	143
5.2 Beschreibung der quantitativ-standardisierten Begleitforschung in Hessen	146
5.2.1 Projektbeschreibung.....	147
5.2.2 Modulbeschreibungen.....	149
5.2.3 Durchführung der Begleitforschung am IQB	151
5.2.4 Stichprobe der Studie 2	152
5.3 Die eingesetzten Instrumente	155
5.3.1 Einstellungen gegenüber den Bildungsstandards.....	156
5.3.2 Kooperation und Zusammenarbeit in der Fachkonferenz.....	157
5.3.3 Kompetenzorientierte Schülertätigkeiten	158
5.3.4 Mathematisches Selbstkonzept der Schülerinnen und Schüler	161
5.3.5 Die Wahrnehmung des Lehrers durch die Schülerinnen und Schüler.....	163
5.3.6 Der standard-basierte Mathematiktest	163
5.3.6.1 Aufgabenauswahl.....	163
5.3.6.2 Testdesign	164
5.4 Vorgehensweise und Analyse von Längsschnittdaten	166
5.4.1 Generierung eines Zeitfaktors	167
5.4.2 Skalenbildung	167
5.4.3 Skalierung von Personenfähigkeiten	168
5.4.4 Umgang mit fehlenden Werten	169
5.4.5 Datenimputation.....	170
5.4.6 Mehrebenenstruktur	174
6 Ergebnisse.....	176
6.1 Drop-Out Analysen.....	176
6.1.1 Drop-Out-Analyse der Studie 1	176
6.1.2 Drop-Out-Analyse der Studie 2.....	181
6.2 Forschungsfrage 1: Merkmalsveränderungen und Entwicklungstendenzen von Mathematiklehrkräften	182
6.2.1 Studie 1 „Erprobung eines Konzeptes zur Implementation der Bildungsstandards in Berliner und Brandenburger Schulen“	183
6.2.1.1 Einstellungen gegenüber der Einführung von nationalen Bildungsstandards	183

6.2.1.2 Zusammenarbeit und Kooperation in der Fachkonferenz.....	191
6.2.1.3 Die mathematischen Kompetenzen.....	198
6.2.1.4 Zusammenfassung der Ergebnisse aus Studie 1	216
6.2.2 Studie 2 „Implementation der Bildungsstandards durch unterstützte Unterrichtsentwicklung“.....	216
6.2.2.1 Einstellungen gegenüber der Einführung von nationalen Bildungsstandards	217
6.2.2.2 Kooperation und Zusammenarbeit in der Fachkonferenz.....	221
6.2.2.3 Der kompetenzorientierte Mathematikunterricht	224
6.2.2.4 Zusammenfassung der Ergebnisse aus Studie 2	235
6.2.3 Zusammenfassung der Forschungsfrage 1	235
6.3 Forschungsfrage 2: Leistungs- und Wahrnehmungsveränderungen von Schülerinnen und Schülern aus einer kompetenzorientierten Perspektive	236
6.3.1 Studie 1 „Erprobung eines Implementationskonzeptes in Berliner und Brandenburger Schulen“:	237
6.3.1.1 Entwicklung des mathematischen Selbstkonzeptes.....	237
6.3.1.2 Entwicklung des mathematischen Interesses.....	239
6.3.1.3 Entwicklung der Lehrerwahrnehmung.....	241
6.3.1.4 Entwicklung der mathematischen Kompetenzen	244
6.3.1.5 Entwicklung der Mathematikleistung	259
6.3.2 Studie 2 „Implementation der Bildungsstandards durch unterstützte Unterrichtsentwicklung“ – Ergebnisse aus Schülersicht	263
6.3.3 Zusammenfassung der Forschungsfrage 2	273
7 Diskussion und Ausblick	275
7.1 Veränderung des Unterrichtsangebots	275
7.2 Veränderungen der Schülerwahrnehmung und der Mathematikleistung.....	283
7.3 Grenzen bei der Messung von Unterrichtsmerkmalen	286
7.4 Ausblick	287
8 Literatur.....	290
9 Abbildungsverzeichnis	316
10 Tabellenverzeichnis	319
11 Anhang	323

Abkürzungsverzeichnis (in alphabetischer Reihenfolge)

BLK	Bund-Länder-Kommission
CBAM	Concern Based Adoption Model
CHiK	Projekt „Chemie im Kontext“
COACTIV	„ <u>Cognitive Activation</u> in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics“ (deutscher Titel: „Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz“)
DESI	Deutsch Englisch Schülerleistungen International
FOR.MAT	"Bereitstellung von Fortbildungskonzeptionen und -materialien zur kompetenz- bzw. standardbasierten Unterrichtsentwicklung"
ICC	Intraclass Correlation Coefficient oder Intraclass correlation
IQB	Institut zur Qualitätsentwicklung im Bildungswesen
IRT	Item Response Theory
KMK	Ständige Konferenz der Kultusminister der Länder
KOU	Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften
MARKUS	M athematik - Gesamterhebung R heinland-Pfalz: K ompetenzen, U nterrichtsmerkmale, S chulkontext
NCTM	National Council of Teachers of Mathematics
PALMA	Projekt zur Analyse der Leistungsentwicklung in Mathematik
PISA	Program for International Student Assessment
PLG	Professionelle Lerngemeinschaft
PROSA	P roblemlösen und S elbstregulation fördern - A usbildungsprogramm
QuaSUM	Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik
SINUS	Programm zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“
TIMSS	Trends in Mathematics and Science Study
VERA	„ <i>VER</i> gleichs <i>Ar</i> beiten“

1 Einleitung

In dieser Arbeit werden die Ergebnisse der quantitativen Begleitforschung zur Implementation der Bildungsstandards vorgestellt, die in der Zeit von 2007 bis 2009 generiert wurden. Ausgehend von den 2003/2004 beschlossenen und verbindlich eingeführten Bildungsstandards werden relevante Merkmale in den Blick genommen, die den Veränderungs- und Entwicklungsprozess von Personen abbilden, die durch eine unterstützte Unterrichtsentwicklungsmaßnahme diese Bildungsinnovation zu implementieren versuchen. Das Ziel dieser Untersuchung ist die Darstellung eines Entwicklungsprozesses auf Schulebene, der Auskunft darüber gibt, inwieweit es Lehrkräften gelingt, ihr Verhalten in Schule und Unterricht dahingehend zu ändern, einen auf Kompetenzerwerb ausgelegten Unterricht zu gestalten.

Aus den Erfordernissen internationaler, gesellschaftlicher, wirtschaftlicher und auch politischer Entwicklungen entwickelte sich auch in Deutschland die Erkenntnis, die Form des Vorbereitens junger Menschen auf eine verantwortungsvolle Teilhabe am gesellschaftlichen Leben zu überdenken. Die Definition des Outcomes von Schule erfuhr eine Neugestaltung hinsichtlich der für die Fortentwicklung einer Gesellschaft nützlichen Kombination von Fähigkeiten und Fertigkeiten des Humankapitals. Ohne die Teilnahme an internationalen Leistungsvergleichen wären die Art und der Ausprägungsgrad von Fähigkeiten und Fertigkeiten deutscher Jugendlicher nicht in dem Ausmaß sichtbar geworden, wie sie zu der aktuellen Situation im Bildungswesen führten. Angeregt durch Diskussionen über die Wettbewerbsfähigkeit deutscher Schulabsolventen entschied sich Deutschland für eine Umsteuerung in der Bildungspolitik, um den Bedürfnissen nach kompetenten Problemlösern gerecht zu werden. Dadurch entwickelte sich eine neue Fokussierung auf den Lernort Schule, in dessen Verantwortung mehrheitlich die Ausbildung von Kompetenzen liegt. Mit Blick in andere Länder, die den Erfordernissen der internationalen Vergleichsstudien entsprachen, stellten deutsche Entscheidungsträger ein grundlegendes Defizit fest, das die Grundlage für Zusammenhangs- und Wirkungsanalysen bei politischen Entscheidungen darstellt. So folgte ein bildungspolitisches Maßnahmenpaket, das zu einer generellen Umsteuerung der Sicht auf Lehren und Lernen führen soll. Der Beschluss der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) zur Einführung und verbindlichen Umsetzung von einheitlichen länderübergreifenden Bildungsstandards wird begleitet von weiteren Erfolg versprechenden Maß-

nahmen auf dem Weg zu einem neuen Verständnis von Schule¹ (KMK, 2003a). Auf der Basis von Bildungsstandards soll es durch Monitoringstudien und landesweiten Lernstandserhebungen gelingen, Daten zu generieren, die die Grundlage für Steuerungswissen und bildungspolitische Entscheidungen bilden. Eingebettet in eine Gesamtstrategie zur Qualitätssicherung im Bildungswesen, die von der KMK im Jahre 2006 beschlossen wurde, werden in Zukunft Daten sowohl auf Systemebene als auch auf Schulebene generiert, die einerseits zum Aufbau eines Datenpools dienen und andererseits die erreichten Fortschritte mit Blick auf die internationale Wettbewerbsfähigkeit überprüfen. So erfolgt in Zukunft eine Orientierung am Output des Schulsystems, was eine Messung der Leistungen von Schülerinnen und Schülern zu bestimmten Zeitpunkten in ihrer Schullaufbahn impliziert. Dieser definierte Lernoutput wird in den national geltenden Bildungsstandards in Form von Kompetenzen beschrieben und unterscheidet sich hinsichtlich der bisher vorherrschenden Inputorientierung dahingehend, dass ein Vergleich von Curriculumsinhalten und Lernmaterialien in den Hintergrund tritt und der Fokus auf die Erfassung von Basiskompetenzen in variierenden Problemsituationen gelegt wird (Prenzel et al., 2005). Damit wird eine Paradigmenwende eingeleitet, innerhalb der das Element Bildungsstandards den Referenzrahmen für qualitätssichernde Maßnahmen wie Schulevaluationen, Vergleichsarbeiten und Monitoringstudien bildet (KMK, 2006)². Der Innovationsgehalt von Bildungsstandards hängt während dieser Umsteuerung von der Intensität der Kommunikation auf den verschiedenen Ebenen des Bildungssystems ab, d.h. je intensiver die Intention, Funktion und der Inhalt der Bildungsstandards diskutiert wird, desto vertrauter und handhabbarer wird dieses neue Element im System Schule und desto wahrscheinlicher wird eine erfolgreiche Implementation (Oelkers & Reusser, 2008). Eine handlungswirksame Einführung von Bildungsstandards gilt dann als erfolgreich, wenn Lehrkräfte einen Unterricht initiieren, der die in den Standards formulierten Kompetenzen aufbaut und fördert. Klieme et al. (2003) machen deutlich, dass neben dem politischen Willen und der Zwecksetzung besonders der professionelle und persönliche Vorteil bei der Arbeit mit den Bildungsstandards kommuniziert werden müsse. Nur eine Akzeptanz, die durch Information, Handreichungen und Fortbildungen aufgebaut werden soll, erhöht die Erfolgswahrscheinlichkeit einer handlungswirksamen Umsetzung von Bildungsstandards (Klieme et al., 2003; Oelkers & Reusser, 2008). Diese handlungswirksame Umsetzung

¹ Als weitere Maßnahmen werden hier die erweiterte Autonomie von Schulen, die Schulwahlfreiheit der Eltern oder auch die Reformbemühungen in der Lehreraus- und fortbildung betrachtet.

² Bethge und Priebe (2008) beschreiben die Paradigmenwende als eine „Orchestrierung“ schulpädagogischer und bildungspolitischer Maßnahmen.

der Bildungsstandards zielt auf eine Umorientierung auf der Ebene des Unterrichts, dessen Durchführung vom Prinzip der Kompetenzorientierung her gedacht werden soll. Jede Stunde sollte sich folglich nach der Zielformulierung evaluieren lassen: Welche Kompetenzen konnten heute aufgebaut werden? (Blum, 2006b; Ziener, 2006). Inwieweit dieses Umdenken für Lehrkräfte anschlussfähig an ihr bestehendes Wissen oder hochinterpretativ ist, hängt aus kognitiver Perspektive mit ihren bestehenden Wissensstrukturen und ihrem darauf aufbauendem Lehrverständnis zusammen (Spillane, Reiser & Reimer, 2002). In diesem Zusammenhang weist Fullan (2001) auch auf ein auftretendes Verständnisproblem bei Lehrkräften hin. Diese könnten im Sinne eines „false clarity“ eine mit den Bildungsstandards implizierte Unterrichtskonzeption als nur eine geringe Abweichung dessen betrachten, was sie immer schon praktizierten und damit den geschaffenen Freiraum für Unterrichtsgestaltung unterschätzen, den Bildungsstandards bieten. Tatsächlich sollte ein kompetenzorientierter Unterricht auf die in den Bildungsstandards formulierten Kompetenzen fokussieren (Helmke, 2009). In der Reflexion vergangener Vergleichsstudien konnte für den Mathematikunterricht beobachtet werden, dass deutsche Schülerinnen und Schüler erfolgreich technische Routinen abarbeiten und kleinschrittige Aufgabenstellungen bearbeiten können; Defizite sind jedoch in den Bereichen des mathematischen Modellierens, Argumentierens und Kommunizierens beobachtbar (Baumert et al, 2000a). Ein standard-orientierter Unterricht soll diese Defizite verringern und anschlussfähige Kompetenzen ausbilden. Wenngleich die Literatur aus den angloamerikanischen Ländern darauf aufmerksam macht, dass die Implementation einer bildungspolitischen Innovation mehr als zehn Jahre in Anspruch nehmen kann (Fullan, 2007; Oelkers & Reusser, 2008), möchte diese Arbeit einen Einblick in beginnende Entwicklungsprozesse und Umsetzungsversuche der Lehrerschaft geben und einen Beitrag zur Implementationsforschung im deutschsprachigen Raum leisten.

1.1 Problemaufriss

Ausgehend von dem Beschluss der KMK über die verbindliche Einführung länderübergreifender Bildungsstandards sehen sich die Akteure im Bildungssystem mit einer Innovation konfrontiert, die es zu implementieren gilt (KMK, 2003a). Strategisch wurde damit *top-down* eine verbindlich einzuführende Maßnahme verabschiedet, die sich aus dem politischen Bestreben ergab, Schul- und Unterrichtsentwicklung an die Erfordernisse eines in Kompetenzstrukturen denkenden, transparenten und rechenschaftspflichtigen Bildungssystems auszurichten. Wenngleich sich ein Accountability-System in Deutschland

noch im Aufbau befindet, bilden die Bildungsstandards dennoch den Referenzrahmen für ein aus Assessment, Vergleichsarbeiten und Unterrichtsentwicklung bestehendes Monitoringsystem (KMK, 2006; Köller, 2008; Oelkers & Reusser, 2008), dessen Erkenntnisse nach dem Prinzip eines spiralförmigen Zyklus in das System zurückgespielt werden und somit fortschreitende und andauernde Entwicklungsprozesse initiieren. Der Erfolg einer solchen Reform hängt von der Implementation der im qualitätssichernden Strategiemodell verankerten Komponenten ab. Für das Instrument „Bildungsstandards“ als Referenzrahmen zur Messung von Schuloutput wird die Implementation an ihren wichtigsten Wirkungsorten, den Schulen und Klassenzimmern, entscheidend sein (Oelkers & Reusser, 2008). Die gemessenen Schülerleistungen stehen nunmehr stellvertretend für das Ergebnis von Schulentwicklungsprozessen und Unterrichtsqualität. Diese Rückkopplung erfolgt inferent durch die Analyse von Schülerkompetenzen innerhalb eines Schulkontextes. Dabei wird davon ausgegangen, dass der Schüleroutput in Form von Kompetenzen Rückschlüsse auf die im Unterricht initiierten Lerngelegenheiten bildet (Ditton, 2000; Helmke, 2009; Maak Merki & Schwippert, 2008). Umgekehrt wird durch das Konzept der Kompetenzorientierung ein Unterrichtsprinzip propagiert, das die Erfüllung der mit den Standards implizierten Anforderungen verspricht. Ein solcher kompetenzorientierter Unterricht bedeutet für das Fach Mathematik eine Fokussierung auf die im Kompetenzmodell festgelegten allgemeinen mathematischen Kompetenzen; argumentieren, kommunizieren, modellieren, Probleme lösen, Darstellungen verwenden und technisch arbeiten (Abschnitt 2.2.1). Die Analyse bisheriger Unterrichtseinheiten im Fach Mathematik deutet auf eine primär lehrerzentrierte Vermittlung von Wissen hin, die jedoch den aktiven Erwerb von Kompetenzen durch die Schülerinnen und Schüler bereits in dieser Vermittlungsphase des Zeigens, Darstellens, Erklärens und Begründens durch den Lehrer außer Acht lässt (Lersch, 2010). Eine Verlagerung dieser Unterrichtsgestaltung hin zu einer schülerzentrierten kognitiven Aktivierung verspricht die Ausbildung der geforderten Problemlösekompetenzen. Wenngleich Schott und Ghanbari (2008) auf die Notwendigkeit von Grundlagenforschung über die Struktur, Dimensionalität und die Ausprägung der Kompetenzen hinweisen, hat die Reform „Bildungsstandards“ ihre Arbeit aufgenommen. Mit Verweis auf Klieme und Leutner (2006) wird deutlich, dass es „bei der Formulierung von theoretischen Modellen zur Beschreibung und Erklärung von Kompetenzen zum einen um die Definition von Strukturen, zum anderen um die Charakterisierung von Niveaus [geht].“ (Schott & Ghanbari, 2008, S. 20). Für die Erfassung und Definition von Kompetenzen lässt sich somit eine unter Zeitdruck vorangetriebene Entwicklung beobachten, in

der Kompetenzmodelle und Kompetenzstufenmodelle gleichzeitig entwickelt und wirksam werden³.

Eine andere Problematik wird sichtbar, wenn man sich parallel zur wissenschaftlichen Entwicklung von Messinstrumenten die Bemühungen in der Schulpraxis ansieht. Die Einführung und Umsetzung der Bildungsstandards wird durch die Entwicklung von kompetenzorientierten Unterrichtsmaterialien, verschiedenen Projekten zur Unterrichtsentwicklung und Lehrerfortbildungen vorangetrieben. Allein in Deutschland sind damit 800.000 Lehrpersonen angesprochen, denen die Idee und Intention von Bildungsstandards vermittelt werden muss und die das Konzept eines kompetenzorientierten Unterrichts verinnerlichen sollen (Oelkers & Reusser, 2008). Didaktiker und Erziehungswissenschaftler haben zwar eine Vorstellung und ein allgemeines Konzept eines solchen Unterrichts, der Vorteil und die Wirkung eines auf Kompetenzen ausgerichteten Unterrichts sind bisher jedoch noch nicht gesichert. Wie ein kompetenzorientierter und standardbasierter Unterricht im Kontext der jeweiligen Schule aussehen soll, ohne in die kritisierte Praktik des „teaching to the test“ zu verfallen, kann nur mit Längsschnittstudien untersucht werden.

Im Rahmen dieser Arbeit sollen die komplexen Prozesse untersucht werden, die sich vorrangig mit der Analyse von Entwicklungs- und Veränderungsprozessen in den Einstellungen und der Unterrichtsgestaltung von Lehrkräften beschäftigen. Dazu wird im Theorie-Teil zunächst der Frage nachgegangen, wie sich der Unterricht ändern soll. Es werden Zielvorstellungen auf der Grundlage empirischer Studien formuliert, deren Konzept die Merkmale eines qualitativ hochwertigen Mathematikunterrichts beinhaltet. Es wird das Bild eines Unterrichts gezeichnet, in dem die Überwindung des routinierten Abarbeitens von Rechenfolgen möglich ist und die Verstehenskomponente eines zusammenhängenden mathematischen Konzeptwissens im Vordergrund steht. Die internationalen Untersuchungen im Rahmen der TIMS-Studie liefern hierfür ertragreiche Impulse. Am Ende des Gliederungspunktes 2 soll die Vision eines standardorientierten Mathematikunterrichts entstanden sein, der die in den Bildungsstandards formulierten Kompetenzen aufzubauen vermag.

Im Anschluss wird überlegt, wie ein solcher Unterricht in die Schulpraxis implementiert werden kann. Dazu erfolgt ein Blick in bereits bestehende Theorien zur Implementation einer schulischen Innovation. Die sehr komplexen Theorien entstanden in der Folge der

³ Zur Unterscheidung von Kompetenzmodellen und Kompetenzstufenmodellen vgl. die Abschnitte 2.2.1 und 2.2.2.

Implementationsbemühungen vergangener Schulreformen in den USA. Sie bieten eine Fülle von Erklärungsansätzen zur Steuerung eines Implementationsprozesses. Dennoch verbleibt aus der Implementationsforschung oftmals die ernüchternde Feststellung, dass der Zeitraum bis zur vollständigen Übernahme einer Innovation sehr lang sein und eine Vielzahl von Faktoren das intendierte Ergebnis erheblich beeinflussen kann. Konzeptionell scheint die Arbeit in professionellen Lerngemeinschaften ein wesentliches Erfolgskriterium zu sein. Aufbauend auf dieser Erkenntnis wird im Anschluss ein Konzept zur Implementation der Bildungsstandards vorgestellt, dass die wesentlichen Merkmale gelingender Unterrichtsveränderungen aufgreift und dessen Wirkung auf die Einstellungen und Unterrichtsgestaltung der Lehrkräfte untersucht wird.

In Studie 1 erhalten daher interessierte Projektgruppen des Faches Mathematik (Fachkonferenzen) die Gelegenheit, ihren Unterricht mit Hilfe einer externen Setkoordinatorin, die durch inhaltliche und didaktische Impulse die Arbeit in den Fachkonferenzen unterstützen soll, dahingehend zu verändern, dass die mathematischen Kompetenzen der Schülerinnen und Schüler im Sinne der Bildungsstandards aufgebaut werden. In diesem zweijährigen Projekt bot sich den Lehrkräften mehrmals pro Schulhalbjahr die Gelegenheit, neue Unterrichtsmaterialien zu entwickeln, zu erproben und in der Fachgruppe zu besprechen. Die Arbeit wurde fortlaufend wissenschaftlich begleitet und durch wissenschaftlichen Input bereichert. Parallele Entwicklungen im Bundesland Hessen boten die Gelegenheit, das im Institut zur Qualitätsentwicklung im Bildungswesen (IQB) entwickelte Konzept mit dem SINUS-Fortbildungskonzept (Programm zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“) zu vergleichen, weshalb die Studie 2 einen Vergleich mit hessischen Lehrkräften ermöglicht.

Eine zweite forschungsrelevante Frage wird untersuchen, ob die indirekte Wirkungskette eines kompetenzorientierten Unterrichts Auswirkungen auf die Unterrichtswahrnehmung und die mathematischen Leistungen der Schülerinnen und Schüler hat. Hierzu wurden in der wissenschaftlichen Begleitung mehrmals pro Schuljahr Testungen auf der Grundlage kompetenzorientierter Mathematikaufgaben durchgeführt. Die Untersuchung der Schülerleistung evaluiert in pragmatischer Weise die Einführung der Bildungsstandards im Hinblick auf die national und international geforderten Kompetenzen des Modellierens.

Die Ergebnisse werden im Anschluss kritisch diskutiert und in den Rahmen paralleler Entwicklungen im Zeitraum des Entstehens dieser Arbeit gesetzt.

1.2 Eingrenzung des Untersuchungsgegenstandes

Bildungsstandards werden von Klieme et al. (2003) und Oelkers und Reusser (2008) als Instrument, Referenzrahmen, Reformelement, Komponente in einem Steuerungsmodell und Oberbegriff für Kompetenzbeschreibungen bezeichnet. Aufgrund der maßgeblich starken Gewichtung der Definition von Standards in der Klieme-Expertise führen sie zu einer Arbeitsdefinition, die sie als Begriff für domänenspezifische Kompetenzzuschreibungen ausweisen. Nach Lersch (2010) ist das Neue und Besondere an den Standards die Tatsache, dass sie die Erwartungen an die Ergebnisse schulischer Lehr-Lern-Prozesse in Form von Kompetenzen beschreiben. Es handelt sich um eine neue Idee, eine andere Betrachtungsweise, die andere Maßnahmen impliziert, um ihr zu begegnen. Zieht man die Definition von Rogers (2003, S. 12) über eine Innovation heran, die er als eine von Individuen als neu wahrgenommene Idee, Praktik oder Objekt beschreibt, dann handelt es sich bei der inhaltlichen Ausgestaltung der Bildungsstandards um eine Innovation. Holtappels (2005) führt eine Trennung der Bedeutung von Innovationen und Schulreformen durch, indem er Innovationen mit dem Vorgang der Einführung von bestimmten Neuerungen assoziiert und unter dem Reformbegriff die Umformung von Bestehendem meint. Beides vereint er in dem Begriff Schulentwicklung, der nicht nur auf die Entwicklung und Konzipierung von Reformmodellen hindeutet, sondern zusätzlich die Adaptierung konkreter Innovationen in und von der Praxis impliziert. Einzelschulische Entwicklung wird damit zum Kern einer standardbasierten Schulreform, deren Ergebnisse im Sinne eines technokratischen Reformansatzes nach Herzog (2010) in einem zirkulären Kreislauf wieder in die Systemebene rückgefüttert werden. Fullan (2007) unterscheidet diesbezüglich zwei entscheidende Begrifflichkeiten im Bereich der schulischen Innovationsforschung: *innovation* und *innovativeness*. Während die Innovation ein bestimmtes Programm oder Produkt anspricht, dessen Erfolg oder Misserfolg messbar ist, beschreibt die Innovationsfähigkeit die vorhandene Kapazität oder Möglichkeit einer Organisation, sich fortlaufend zu verändern. Bezogen auf die Bildungsstandards bedingen sich beide Begrifflichkeiten gegenseitig. Eine erfolgreiche Implementation bedarf einer schulischen Organisation, die zur Veränderung fähig ist. Aus einer prozessorientierten Perspektive stehen Bildungsstandards am Ende einer Wirkungskette und bilden sozusagen die innovative Zieldimension, von der aus der Kreislauf erneut ansetzt. Rückwärts gedacht ist dem Ziel ein Weg oder ein Prozess vorgelagert, der im hier betrachteten Kontext als Kompetenzerwerb oder –aufbau bezeichnet wird. Dieser Prozess wird für die Schülerinnen und Schüler moderiert durch die Schaffung von geeigneten Lernarrangements. Auf die Institution Schule bezo-

gen wird damit ein kompetenzorientierter Unterricht bezeichnet, der im Sinne Rogers eine Innovation in Form einer neuen Praktik darstellt. Damit wird für die weitere Arbeit das Konzept der Kompetenzorientierung⁴ ebenfalls als eine Innovation betrachtet. Popham (2004, S. 17) beschreibt es so: „Standard-Based Education is nothing more than a posh ends-means model wherein content standards represent intended ends, teaching constitutes the means for achieving those ends, and test results supply the evidence regarding whether the means did, in fact, achieve the intended ends.“.

Ein weiteres für diese Arbeit relevantes Element bilden die Kompetenzmodelle. In dem von Herzog (2010) angewendeten Zweck-Mittel-Schema dienen sie als Scharnier zwischen den Regelkreisen der Politik und der Pädagogik. Da mit ihrer Hilfe kompetenzorientierte Aufgaben zur Überprüfung der Bildungsstandards in einem Modell verortet werden können, sind Kompetenzmodelle als Instrument zu bezeichnen. In den deutschsprachigen Ländern sollen durch Kompetenzmodelle die Prozesse des Kompetenzerwerbs modelliert werden⁵. Sie bilden das für die Bildungsstandards notwendige Kommunikationsmittel, um die komplexen Kompetenzbeschreibungen der Schülerinnen und Schüler in einem Modell verorten zu können. Theoretisch fundierte Kompetenzmodelle befinden sich aktuell in der empirischen Prüfung. Da sie sich in einem fortlaufenden Entwicklungsprozess befinden, sollen sie auf der abstrakten Ebene als Instrument betrachtet werden, nicht aber als Innovation.

Zusammenfassend werden als Untersuchungsgegenstände für diese Arbeit in Anlehnung an Rogers (2003) zwei Innovationen ausgemacht – Bildungsstandards als neue Idee und das Unterrichtskonzept der Kompetenzorientierung als neue Lehrpraktik. Für das weitere Vorgehen wird diese Bestimmung handlungsleitend sein.

Die hohe Komplexität der Innovation Bildungsstandards ergibt sich aus deren multidimensionaler Wirkung auf allen Ebenen des Bildungssystems. Bisher liegen nur wenige Befunde zur Einführung der Innovation Bildungsstandards in Deutschland vor. Wenngleich die Forschung in den USA und Großbritannien bereits auf einen vergleichsweise großen Erfahrungswert zurückgreifen kann, lassen sich in Deutschland kaum Anhaltspunkte finden, wie diese bildungspolitisch initiierte Innovation in den Schulen und Klassenzimmern ankommen und angenommen wird. Mit der Arbeit von Zeitler, Heller und Asbrand (2012) liegt ein konkreter Beitrag zur Implementationsforschung zu den Bil-

⁴ Kompetenzorientierung und standardbasiertes Unterrichten wird gleichbedeutend verwendet.

⁵ Klieme et al. (2003, S. 50): „Indem Bildungsstandards kumulatives Lernen über mehrere Jahrgänge und Niveaustufen hinweg darstellen, verändern sie den Blick auf Unterricht von einer rein fachsystematischen Perspektive hin zu einer stärker schülerorientierten, d.h. an der kognitiven Entwicklung der Lernenden ausgerichteten Perspektive.“

dungsstandards in Deutschland vor. Andere Beiträge aus dem weiteren Umfeld des deutschsprachigen Raums liegen mit den Arbeiten von Oelkers und Reusser (2008) sowie Altrichter und Posch (2007), Böttcher und Dicke (2008), Maag Merki (2004), und Pant (2008a, 2008b) vor. Längsschnittlich generierte Ergebnisse zum Umgang und zur Implementation von Bildungsstandards im Unterricht liegen bislang jedoch noch nicht vor. Mit dem Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA) veröffentlichten Pekrun et al. (2006) zwar eine Längsschnittstudie zur Leistungsentwicklung in Mathematik, der Schwerpunkt des Projektes bezog sich jedoch auf die Untersuchung der Mathematikleistung und des Kontextes von Mathematikunterricht. Holtappels (2005, S. 37) macht darauf aufmerksam, wenn er zur Forschungslinie über Schulentwicklung ausführt, dass prozessbezogene Primärstudien hierzulande rar bleiben. „Sie erfolgen vornehmlich im Zusammenhang mit schulischen Innovationen, vor allem Curriculumsrevisionen, Veränderungen in der schulischen Zeitorganisation oder innovativen Ansätzen im Unterricht.“. Mit der COACTIV-Studie⁶ legten Kunter et al. (2011) eine umfassende längsschnittlich angelegte Untersuchung zur professionellen Kompetenz von Mathematiklehrkräften vor, die es ermöglichte, kausale Wirkrichtungen des Handelns von Lehrkräften auf den Unterricht und die Leistung von Schülerinnen und Schülern zu analysieren. Die Ergebnisse der COACTIV-Studie erlangten große Bedeutung und bestätigten an einer repräsentativen Stichprobe empirisch eine Vielzahl theoretischer Konzeptionen, wie z.B. die Merkmale von Unterrichtsqualität, und eingeleiteter Praxismaßnahmen, wie z.B. das Weiterbildungsprogramm zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (SINUS)“. COACTIV ist damit eine Untersuchung, die Erklärungswissen generiert und ein theoretisches Kompetenzmodell für Mathematiklehrkräfte zur Verfügung stellt. Die Bildungsstandards formulieren dagegen ein Ergebnis, dessen Weg dorthin offen bleibt, jedoch viele Merkmale und Kriterien impliziert. Wenn am Ende der Unterrichtszeit ein bestimmtes Ergebnis herauskommen soll, dann besteht die Notwendigkeit, den Weg dorthin darauf auszurichten, was durch die Kompetenzorientierung im Unterricht beschrieben wird. Diese Arbeit richtet den Blick damit auf das Prozesswissen der Lehrkräfte und die von COACTIV bestätigte Bedeutung und Veränderbarkeit des fachdidaktischen Wissens.

⁶ Vollständiger Name: „Cognitive Activation in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics“.

Unter der Voraussetzung, dass ein fachlicher und/oder fachdidaktischer Input dazu führt, dass Lehrkräfte ihren Unterricht umgestalten, sollen diese Veränderungen abgebildet werden.

2 Kompetenzerwerb auf der Grundlage von Bildungsstandards

Der Anstoß zu den aktuellen Reformbemühungen in Deutschland erfolgte erst durch die öffentlichkeitswirksame Diskussion der ersten PISA-Ergebnisse, wenngleich sich bereits nach der Auswertung der Trends in International Mathematics and Science (TIMS)-Studie ein Handlungsdruck für eine veränderte Schulpolitik offenbarte. Damit folgte Deutschland den Empfehlungen eines Gutachtens infolge der PISA-Ergebnisse (Program for International Student Assessment), in dem den Schulen nach dem Vorbild erfolgreicher PISA-Länder mehr Gestaltungsfreiheit eingeräumt werden soll, gleichzeitig jedoch die Steuerung über Bildungsstandards intensiviert werden soll (Avenarius et al., 2003). Die Bildungsstandards, wie sie im deutschsprachigen Raum konzipiert sind, haben als Performanzstandards den Anspruch, „...Erwartungen (Normen) an fachliches Lernen im Kontext allgemeiner Bildungsziele zu spezifizieren.“ (Maag Merki, 2006, S. 146). Dabei verliert sich der Fokus auf den Input zunehmend und die im Zuge der Standardeinführung gesetzten Ziele geraten referenziell in den Mittelpunkt. Ausdruck erfährt dieser Paradigmenwechsel in den 2006 von der KMK vorgelegten „Plöner Beschlüssen“, die eine Gesamtstrategie zur Qualitätssicherung im allgemeinbildenden Schulsystem beschreiben. Diese aus flächendeckenden Vergleichsarbeiten und nationalen Tests zur Überprüfung der Bildungsstandards (Monitoring), ergänzt um Schulinspektionen und standardorientierter Unterrichtsentwicklung, bestehende zyklische Strategie zur Qualitätsentwicklung, beschreibt in ihrer Logik ein Lernen aus dem Output des vorherigen Messzyklus, dessen Wirksamkeit sich im Folgezyklus messbar zeigt. Doch nicht allein die Überprüfung der Standards führt im Sinne des Reformprozesses zur Steigerung der Unterrichtsqualität, vielmehr bedarf es eines Unterrichts, der Lerngelegenheiten für die in den Standards formulierten Kompetenzen anbietet und damit den eigenaktiven Erwerb von Kompetenzen fördert (Köller, 2010). Standards werden damit zum Zielkriterium in der Unterrichtsplanung und dienen gleichzeitig als Reflexionsmaßstab für den erreichten Lernzuwachs der Schülerinnen und Schüler (Drieschner, 2009). Für die Implementation der Bildungsstandards scheint folglich die Unterrichtsgestaltung auf der Grundlage der Bildungsstandards eine notwendige Maßnahme zu sein, um kompetente Lerner im Sinne einer überprüfbaren Outputsteuerung zu produzieren.

Das folgende Kapitel möchte daher die Konzeption und Formen der Kompetenzorientierung in den Blick nehmen, so wie sie unter der Maßgabe zu beachtender Standards am Ende der Schulzeit sinnvoll erscheinen. Nach der Klärung des Kompetenzbegriffes als Grundlage der weiteren Ausführungen (Abschnitt 2.1), wird konkret das Kompetenzmodell im Fach Mathematik der Sekundarstufe I vorgestellt (Abschnitt 2.2) bevor auf dessen Basis ein kompetenzorientierter Mathematikunterricht beschrieben wird (Abschnitt 2.3).

2.1 Kompetenzen als Ergebnis schulischen Lernens

Die konzeptionellen Grundlagen für die Bildungsstandards wurden von Klieme et al. mit der Expertise „Zur Entwicklung nationaler Bildungsstandards“ (Klieme et al., 2003) vorgelegt. Darin gründen die Autoren ihre Empfehlungen hinsichtlich der Formulierung von Bildungsstandards auf einem Kompetenzbegriff, der fachspezifisch die wesentlichen Ziele pädagogischer Arbeit beschreibt. Konkret resultieren daraus Can do Statements, die die gewünschten Lernergebnisse der Schülerinnen und Schüler als gezeigtes Verhalten zu bestimmten Zeitpunkten in ihrer Bildungsbiographie ausweisen. Im Konsens der Arbeiten mit Bezug auf die Bildungsstandards erfolgt die Anlehnung an einen psychologisch geprägten Kompetenzbegriff, der vordergründig auf empirisch erfassbare Konzepte aus der Expertiseforschung abzielt. Mit der Definition von Kompetenzen nach Weinert (2001) als die bei Individuen verfügbaren oder erlernbaren Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, werden Kompetenzen als hypothetische Konstrukte durch geeignete Operationalisierungen messbar (Köller, 2008). Als bloße Konstrukte der Wissenschaft sind Kompetenzen nicht beobachtbar. „In den Bildungsstandards sind nicht unmittelbar Kompetenzen formuliert, sondern Verhaltensweisen als Indikatoren für die dahinter liegenden, zu erreichenden Kompetenzen beschrieben.“ (Zeitler, Asbrand & Tesch, 2010, S. 24). Wenngleich der Kompetenzbegriff weder neu, noch einheitlich verwendet wird⁷, verspricht die Anlehnung an Weinert (2001) im schulischen Kontext eine über die Fähigkeits- und Wissensfacette hinausgehende Definition, in der ebenso motivationale und handlungsbezogene Merkmale einfließen (Klieme et al., 2003). Der individuellen Ausprägung einer Kompetenz schreibt Weinert den Grad des Zusammenspiels der Facetten Fähigkeit, Wissen, Verstehen, Können, Handeln, Erfahrung und Motivation zu, wobei im Ergebnis den fachbezogenen Fähigkeiten und dem fachbezogenem Wissen zwar eine

⁷ Eine von Weinert (2001) abweichende Kompetenzdefinition im Zusammenhang mit den Bildungsstandards bieten Schott und Ghanbari (2008) im Projekt ComTrans mit ihrer festsetzenden Definition an. Weitere Unterschiede führen Troitschanskaia und Seidel (2011) auf.

starke Gewichtung zufällt, aber eben nicht die alleinige Rolle. Die Berücksichtigung dieser Facetten der Kompetenz führt letztendlich zum Bildungsziel, das durch die formulierten Standards zum Ausdruck kommt. Der Einschluss einer motivationalen Facette macht Kompetenzen zu einem relativ breit gefassten Konstrukt, das kognitionspsychologische und lernpsychologische Elemente vereint (Rupp, Leucht & Hartung, 2006). Der Rückschluss eines gezeigten Verhaltens auf eine vorhandene Kompetenz erfolgt indirekt durch das Lösungsverhalten der Schülerinnen und Schüler bei einer ausreichenden Anzahl vorgelegter Aufgaben, wobei neben ihrer fachspezifischen Kompetenz auch Kompetenzen aus anderen Teilgebieten erschlossen werden. So impliziert das Lösungsmuster von Mathematikaufgaben auch gleichzeitig Lesekompetenzen und Problemlösekompetenzen. Die Abbildung solch komplexer Kompetenzstrukturen führt zu der Freiheit, Aussagen über einzelne Kompetenzen oder ganze Kompetenzbereiche zu ermöglichen, die sich im Fall der Bildungsstandards in Kann-Beschreibungen ausdrücken. Lersch (2010) weist auf eine elementare Eigenschaft von Kompetenzen hin, wie sie im Kontext von Schule zu verstehen sind. Sie sind erlernbar. Damit können Kompetenzen erworben werden und auf bereits vorhandene Kompetenzen kann aufgebaut werden. Kompetenzen beziehen sich auf „Grunddimensionen der Lernentwicklung in einem Gegenstandsbereich“ (Klieme et al., 2003, S. 21). Mit der Betonung des bereichsspezifischen Wissens werden „losgelöste Schlüsselkompetenzen, ohne genetischen Bezug zu einem langfristigen Wissensaufbau und zu Lernbereichen, [...], von Weinert als „inkorrekt“ zurückgewiesen“. (Oelkers & Reusser, 2008, S. 25).

Diese Aspekte von Kompetenz sollen im Folgenden noch etwas genauer beleuchtet werden. Wenn Kompetenzen die grundlegenden Handlungsanforderungen für die Schülerinnen und Schüler in einem Fach darstellen (Klieme et al., 2003), dann soll in einem ersten Schritt die Abgrenzung vom psychologischen Konstrukt der Intelligenz erfolgen, um die im Zusammenhang mit den Bildungsstandards verwendete Kompetenzdefinition besser greifen zu können. Zur weiteren Klärung des Begriffs wird eine Unterscheidung des Begriffspaares Kompetenz und Performanz vorgenommen, da das Konzept der Performanz in der gängigen Literatur häufig als Gegensatz zur Kompetenz genannt wird und die Unterscheidung zwischen einer Handlungsdisposition und einer gezeigten Handlung zur Güte des Kompetenzbegriffs beiträgt. Die Bedeutung des Kompetenzbegriffs für das Bildungswesen soll mit einer dritten Vertiefung hervorgehoben werden, in der die Anschlussfähigkeit des Kompetenzbegriffs an das schulische Lernen herausgearbeitet wird.

2.1.1 Abgrenzung zum Konstrukt der Intelligenz

Mit der Anlehnung an die Kompetenzdefinition von Weinert (2001) im Hinblick auf die Bildungsstandards wird der Begriff der Kompetenz mit der Definition von kontextspezifischen Dispositionen kognitiver Leistungen belegt (Prenzel, Walter & Frey, 2007). Die Kontextualisierung des Kompetenzbegriffs erhält damit die Abhängigkeit von spezifischen Lerngelegenheiten, womit für Prenzel et al. (2007) sowie Baumert et al. (2007) die Abgrenzung vom Intelligenzbegriff erfolgt. Intelligenz, verstanden als Sammelbegriff für allgemeine kognitive Fähigkeiten wird in ihrer Messung eben nicht an solche Lerngelegenheiten gehängt. Die Debatte zur Konkretisierung des Kompetenzbegriffs wurde von Rindermann (2006) ausgelöst, indem er in den Konstrukten der internationalen Schulleistungsvergleichsstudien und den Intelligenztests eine starke Ähnlichkeit und eine enge Beziehung ausmacht. Er führt dazu aus: „Ergebnisse der Schulleistungsstudien sind auf allen Datenebenen in sich höchst homogen, sie korrelieren hoch mit anderen Schulleistungsstudien und mit Intelligenztestresultaten. Die Ergebnisse sind in extremem Maße – weit mehr als aus Intelligenztests selbst bekannt – auf einen gemeinsamen g-Faktor zurückführbar.“ (Rindermann, 2006, S. 84). Dieser g-Faktor steht in der psychometrischen Diskussion für ein Generalfaktormodell zur Überprüfung der Dimensionalität von Tests. Diese Dimensionalität führt zu einem weiteren Unterscheidungsmerkmal zwischen Intelligenz und Kompetenz, das Prenzel et al. (2007) zuerst theoretisch ausarbeiten und im Anschluss empirisch belegen: die erweiterten binnenstrukturierenden Annahmen im Hinblick auf einen funktionalen Kompetenzbegriff. Während die Binnenstrukturierung der Intelligenz an kognitiven Prozessen wie schlussfolgerndes Denken, Gedächtnis, Wahrnehmungsgeschwindigkeit oder räumliches Denken auszumachen ist, erfolgt für die Kompetenzen eine Situierung in Anforderungen, die zur gesellschaftlichen Teilhabe befähigen als auch eine Differenzierung nach kognitiven Prozessen, „..., die zur Bewältigung dieser Anforderungen notwendig sind.“ (Prenzel et al., 2007, S. 129). Die Bereichsspezifität von Kompetenzen wird mit Bezug auf die PISA-Konzeptionen in den unterschiedlichen Konstruktdefinitionen der Erhebungsdomänen deutlich „...durch die Definition von Grundbildung im Hinblick auf Wissensinhalte und Prozesse, die im jeweiligen Bereich erworben werden müssen.“ (Prenzel et al., 2007, S. 129). Die Autoren führen zur empirischen Absicherung die PISA 2003-Daten für den Bereich der Naturwissenschaften an und führen aus, dass die Modellpassung einer der Anzahl der Untertests entsprechenden mehrdimensionalen Skalierung deutlich besser ausfiel, als ein eindimensionales Modell mit der Entsprechung eines Generalfaktormodells. Auch auf der Ebene der Bin-

nendifferenzierung von Kompetenz versus Intelligenz zeigt sich eine bessere Modellpassung für ein mehrdimensionales Modell, in das die Itemdaten der vier PISA-Tests und dem in PISA 2003 eingesetzten kognitiven Fähigkeits- (KFT) - Untertest „Figurale Analogien“ einfließen, gegenüber einem eindimensionalen Modell. Auch Winkelmann, Rotitzsch, Stanat und Köller (2012) können für die mathematische Kompetenz, wie sie auf der Grundlage der Bildungsstandards gemessen wird, zeigen, dass die Dimension der kognitiven Fähigkeit um fachspezifische Anteile ergänzt werden muss, um hinreichend die beobachteten Mathematikleistungen zu erklären. Die Analysen basieren auf Brunners (2006) Resultaten, die eine Präferenz für ein Nested-Faktormodell gegenüber einem g-Faktormodell erkennen lassen, indem die mathematische Kompetenz um eine übergeordnete Kompetenz (z.B. globale Mathematikkompetenz oder allgemeine kognitive Fähigkeit) ergänzt wird. Dieses hierarchische, mehrdimensionale Modell mit einem Generalfaktor und zwei fachspezifischen Faktoren zeigt eine deutlich bessere Modellpassung als ein eindimensionales, nur den g-Faktor berücksichtigendes Modell.

Mit diesem Beleg der Differenzierbarkeit des Kompetenzbegriffs widersprechen Prenzel et al. (2007) und Baumert et al. (2007), Rindermanns Konsequenz der engen Beziehung zwischen Schulleistungsstudien und Intelligenztests. Vielmehr scheint bei Beachtung der Ergebnisse von Helmke und Weinert (1997) aus der Scholastikstudie eine kompensatorische Beziehung zwischen den beiden Konstrukten zu bestehen. Helmke und Weinert (1997) fanden heraus, dass der Einfluss der Intelligenz mit zunehmender mathematischer Kompetenz in der Grundschule abnimmt. Je unvertrauter die Lerninhalte und je weniger Vorwissen vorhanden ist, desto schlechter dient ein Intelligenzfaktor als Vorhersagewert für die mathematische Kompetenz. „Mit steigender Kompetenz nimmt die Vorhersagekraft des Vorwissens in Relation zur Intelligenz systematisch zu.“ (Baumert et al., 2007, S. 124).

Hartig und Klieme (2007) führen zur Binnenstrukturierung von Intelligenz und Kompetenzen aus, dass die Struktur der Intelligenzleistung zu einem wesentlichen Teil durch die Prozesse definiert wird, die für die beobachtbaren Intelligenzleistungen als notwendig und grundlegend betrachtet werden, wohingegen die Strukturen von Kompetenzen und Teilkompetenzen an die Inhalte der interessierenden Situationen, der relevanten Aufgaben und den zur Lösung dieser Aufgaben zu bewältigenden Anforderungen geknüpft sind. Die Autoren führen neben der differenzierten Kompetenzbeschreibung für Mathematik in PISA 2003 die Deutsch Englisch Schülerleistungen International (DESI) -Studie an, in

der die Kompetenzstruktur aus den Anforderungen spezifischer kommunikativer Situationen und Handlungen abgeleitet wird.

In dem ausdrücklichen Bezug der Kompetenzen zu bestimmten Kontexten sehen Hartig und Klieme (2006) weiterhin das Unterscheidungsmerkmal der Erlernbarkeit von Kompetenzen begründet. Diese Eigenschaft ergibt sich zwangsläufig aus der inhaltlichen Ausgestaltung des Kompetenzbegriffes. So ist ein Kompetenzerwerb und –aufbau gewünscht und messbar, wodurch sich auch erklären lässt, dass für den Kompetenzbegriff ein interpretierbarer Nullwert möglich ist. Kontextspezifische Kompetenzen setzen in den jeweiligen Kontexten Erfahrungen voraus. Ist ein Lerner oder eine Lernerin niemals mit einer solchen Situation konfrontiert worden, kann die entsprechende Kompetenz fehlen. Hingegen ist ein Nullwert als interpretierbarer Intelligenzwert nicht anzunehmen, da jede gesunde Person zu einem gewissen, messbaren Grad über bestimmte Grundfähigkeiten verfügen sollte. Die Unabhängigkeit des Intelligenzwertes vom Kontext lässt jedoch einen Grad an Generalisierbarkeit der damit beschriebenen Leistungsdispositionen zu, den die Kompetenzaussagen gerade durch die Kontextgebundenheit nicht erreichen können. Im Hinblick auf diese Erlernbarkeit und damit auch Veränderbarkeit von Kompetenzen argumentiert Rindermann (2007), dass Intelligenz ebenso veränderbar sei und zieht dazu die von Klauer (2001) durchgeführten Trainings zur Entwicklung kognitiver Grundfähigkeiten heran. Rindermann (2006) verweist damit auf lediglich graduelle Unterschiede der Veränderbarkeit zwischen Kompetenzen und Intelligenz. In Teilen wird diese Aussage von Köller und Karim (2001) sowie Baumert, Trautwein und Artelt (2003) bestätigt. Sowohl aus den BIJU (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter) - als auch den PISA 2000-Daten konnte gezeigt werden, dass differenzielle Lernmilieus, unter Beachtung der Schulform, nicht nur für die Entwicklung der Schulleistung, sondern auch für die Entwicklung der Intelligenz“ eine Bedeutung haben, wobei die Fachleistungen sensibler reagieren als die Intelligenz. „Intelligent angelegte und kognitiv herausfordernde Lerngelegenheiten im Fachunterricht der Schule fördern primär den Erwerb fachspezifischen Wissens, aber auch Intelligenz. Ein Training des schlussfolgernden Denkens aber – so nützlich dies auch sein kann – ersetzt weder den Mathematik- noch den Sprachunterricht.“ (Baumert et al., 2007, S. 125).

Wenngleich zwischen Intelligenz und Kompetenz klare konzeptionelle Unterschiede auszumachen sind, so lassen sich die Konstrukte empirisch nicht ganz so klar trennen. Leutner, Klieme, Meyer und Wirth (2004) berichten Korrelationen zwischen Intelligenz-

und Kompetenzmaßen aus PISA 2003 von $r = .74$ (messfehlerbereinigt), wobei die Wirkrichtung zwischen den Leistungsvariablen aufgrund der Querschnittsdaten nicht auszumachen ist. Intelligenz als reine kognitive Fähigkeit weist allerdings bereits per Definition eine Schnittmenge mit den Kompetenzen auf, wenn diese als kognitive Leistungsdispositionen beschrieben werden. Empirisch lässt sich die Integration der Intelligenz in das Kompetenzkonstrukt durch Brunner (2006) belegen. In seinem Nested-Faktormodell beeinflusst der allgemeine Intelligenzfaktor g die Lösung aller Aufgaben (Itemdaten), während die mathematikspezifische Fähigkeit nur auf die Performanz der Subskalen des PISA-Mathematiktests lädt (Baumert et al., 2007). Als Performanz werden die beobachteten Daten von Tests bezeichnet, womit sich ein weiteres Feld der Kompetenzdefinition öffnet (Abschnitt 2.1.2). Brunner und Krauss (2010, S. 121) machen mit dem Nested-Faktormodell deutlich, wie wichtig die Konzeption einer fachspezifischen Kompetenz ist und schlussfolgert, dass es für eine eindeutige Interpretation einer Mathematikkompetenz notwendig ist, „...Mathematikaufgaben zu entwickeln, die sich nur unter Einsatz einer mathematikspezifischen Fähigkeit lösen lassen, die explizit nichts mit schlussfolgerndem Denken [Anm. Intelligenz wird in der Arbeit von Brunner (2006) durch schlussfolgerndes Denken repräsentiert] zu tun hat (Baumert, Lüdtke, Trautwein & Brunner, 2009)“. Mit ihren Analysen konnten Brunner (2006) und Baumert et al. (2009) zeigen, dass Schülerleistungen nicht nur auf einen reinen Intelligenzfaktor zurückzuführen sind, sondern auch fachspezifische Komponenten aufweisen, die durch geeignete Lerngelegenheiten beeinflusst werden können. Für die Bildungsstandards konnte diese Mehrdimensionalität im Fach Deutsch von Bremerich-Vos, Böhme und Robitzsch (2009) und im Fach Mathematik von Winkelmann und Robitzsch (2009) gezeigt werden. Die Leistungen in beiden Fächern werden als mehrdimensionale Konstrukte identifiziert, die neben einem Intelligenzmaß fachspezifische Anteile aufweisen. In Anlehnung an Weinert (2001) wurde die Kompetenzdefinition für die Bildungsstandards um motivationale, volitionale und soziale Aspekte erweitert. Köller (2008) führt dazu aus, dass die Operationalisierung der Teilfacetten eines Kompetenzbereiches eine valide Messung jeder Facette erlaubt und zusätzlich zu den Aufgaben zur Überprüfung von fachspezifischen Kompetenzen Fragebögen eingesetzt werden, die motivationale Aspekte erfassen. Zu diesem Vorgehen entschied man sich bei der Anwendung eines breiten Kompetenzbegriffes bei der Messung der naturwissenschaftlichen Kompetenz in PISA.

Im nachfolgenden Abschnitt soll die Problematik der Testperformanz näher beleuchtet werden. Der Zusammenhang mit dem Kompetenz- und Intelligenzbegriff wird deutlich, wenn Zimmer et al. (2007, S. 196) in den vertiefenden Untersuchungen zur PISA-Studie ausführen: „An dieser Stelle ist zu betonen, dass wir uns auf die Performanz der Jugendlichen in den Domänen des PISA-Tests und nicht auf ihr generelles Leistungspotential (z.B. im Sinne von „Intelligenz“) beziehen. Daher bezeichnen wir die von uns identifizierten Gruppen als besonders leistungsfähige oder *hochkompetente* Schülerinnen und Schüler.“ (Hervorhebungen im Original). Im Gegensatz zur Konstruktabgrenzung beziehen sich die folgenden Ausführungen weniger auf die Validität der Kompetenzmessung, sondern fokussieren eher auf die Reliabilität (im weitesten Sinne), d.h. vielmehr gerät das gezeigte Verhalten von Personen und die darauf beruhenden Implikationen in den Fokus.

2.1.2 Performanz und Kompetenz

Die Gegenüberstellung der Begriffe Performanz und Kompetenz ist in den Sozialwissenschaften im weiteren Sinne eine häufige Methode, um die erlernten Fähigkeiten und Fertigkeiten von den tatsächlich in der Diagnosesituation gezeigten Handlungen zu unterscheiden. Folgt man Chomsky (1973), dann wird diese Unterscheidung in dem Sinne bedeutend, dass die Kompetenz prinzipiell als Handlungsdisposition verstanden wird, während die Performanz durch das tatsächliche Handeln in einer bestimmten Situation gekennzeichnet ist (Gillen & Kaufhold, 2005). Die Autoren kommen zu dem Schluss, dass „...Kompetenzen aufgrund ihres Potentialcharakters nicht vollständig erfassbar sind, da es nicht möglich ist, eine umfassende objektive Erkenntnis über subjektive Variablen zu erlangen.“ (Gillen & Kaufhold, 2005, S. 367). Auch Ziener (2006) bezieht sich in seiner Verständnisunterscheidung zwischen Kompetenz und Performanz auf Chomsky, indem er für den Bereich der Sprachkompetenz schlussfolgert, dass ein Sprechakt erst zustande kommt, wenn Sprachkompetenz vorhanden ist. Mit steigender Kompetenz impliziert er ein größeres Wahlrepertoire an Sprechmöglichkeiten, die eine Person aktivieren und realisieren kann, um zu kommunizieren. Kompetenz ist im Verständnis von Ziener (2006) der Performanz vorgelagert und bildet die Voraussetzung zur Handlungsfähigkeit. Schott und Ghanbari (2008) beschreiben die Beziehung zwischen Kompetenz und Performanz dahingehend, dass eine Kompetenz immer nur indirekt über die entsprechende Performanz diagnostiziert werden kann. Die Beziehung der beiden Begriffe zueinander ist durch das Merkmal der Beobachtung gekennzeichnet, denn nur eine beobachtete Handlung einer Person lässt den Schluss auf die Art und Ausprägung einer bestimmten Fähigkeit zu.

Die Problematik des „Könnens“ und „Wollens“ wird durch eine Vielzahl anderer Faktoren beeinflusst und kann die Ergebnisse von Vergleichsstudien empfindlich verzerren. Während Kiplinger, Haug und Ademi (2000) den Zusammenhang von Mathematikleistung und Sprachniveau bei sprachlastigen Mathematikaufgaben bestätigen, werden weitere Erklärungsfaktoren für eine schlechte Testperformanz in der Instruktion, der Fähigkeit zu dekodieren, auf Texte zu antworten und zu vergleichen gesehen (Hanson et al., 1998). Damit kommt der Performanz eine essentielle Rolle für die Kompetenzdiagnose zu, die unter unglücklichen Umständen zu falschen Implikationen führt.

Aus psychologischer Sicht wird die Kompetenz damit als latente Fähigkeit bezeichnet, die bei Klieme und Hartig (2008) eher keine Frage der Belegung bestimmter Dispositionen ist, sondern vielmehr eine Frage der mentalen Bedingungen, die dem erfolgreichen Handeln zugrunde liegen. Im psychologischen Sinne geht es bei der Kompetenzdiagnostik um die Modellierung dieser mentalen Bedingungen, wobei je nach fachspezifischer Eingrenzung unterschiedliche Konstrukte herangezogen werden. Schott und Ghanbari (2008) führen jedoch aus, dass es einen Unterschied macht, wie man eine Kompetenz beschreibt und welche psychologischen Prozesse man modelliert.

Köller (2008, S. 165) führt für die als can-do-Statements formulierten Standards aus: „Die Standards sind hier als erfolgreich gezeigtes Verhalten definiert und man unterliegt einem Missverständnis, wenn man dieses Verhalten (die Performanz) mit Kompetenz gleichsetzt.“ Als Performanz wird damit die konkrete Ausführung einer Handlung verstanden, während die Kompetenz als latente Fähigkeit über definierte Indikatoren oder Situationsbeschreibungen gemessen werden kann (Köller, 2008; Schott & Ghanbari, 2008). Schecker (2012, S. 229) formuliert ganz klar: „Competence tests first of all test performance.“. Er fügt jedoch hinzu, dass die Performanz nahezu immer geknüpft ist an anwendbares Wissen oder anwendbare Fähigkeiten in einem bestimmten Bereich. Die Fähigkeit, ein Experiment erfolgreich durchzuführen kann beispielsweise einerseits auf eine allgemeine Experimentalkompetenz zurückgeführt werden oder andererseits auf eine Fähigkeit, die eng mit dem speziellen Experiment verknüpft ist. Werden jedoch weitere Lerngelegenheiten konstruiert, in denen der Lerner oder die Lernerin die allgemeinen Experimentalkompetenzen anwenden muss, so wird die Diagnose einer allgemeinen Experimentalkompetenz wahrscheinlicher. Solche häufig gestellten Performanzsituationen sollen es den Schülerinnen und Schülern ermöglichen, eine Kompetenz aktiv aufzubauen und das erlangte Wissen unmittelbar anzuwenden.

In dieser didaktischen Perspektive wird die Unterscheidung zwischen Performanz und Kompetenz in der Methodik eines kompetenzfördernden Unterrichts deutlich. Während die Kompetenz das Produkt einer Wirkungskette darstellt, stellt sich die Performanz als Prozessvariable dar, indem eine Performanzsituation nach Lersch (2010) als Lernmethode verstanden wird (Abbildung 1).

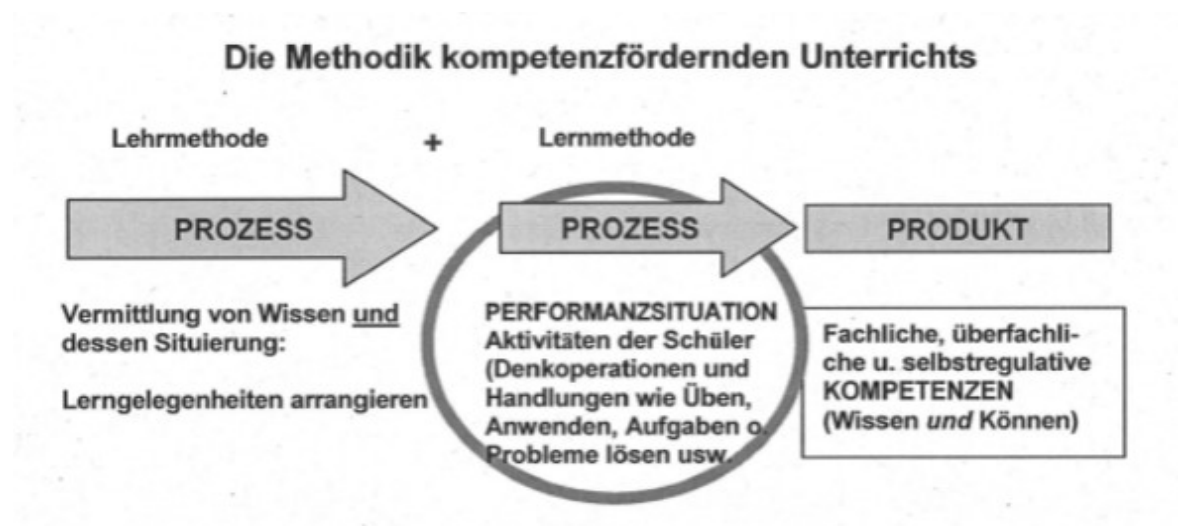


Abbildung 1. Kompetenz und Performanz aus didaktischer Perspektive nach Lersch, 2010.

Mit Bezug auf Klingberg (1986) führt Lersch (2010, S. 7) aus, dass es für den Erwerb von Kompetenzen gilt, „...die Schülerinnen und Schüler permanent in solche möglichst selbstständig zu bewältigende Performanzsituationen zu verwickeln, ...“. Die Performanz im didaktischen Sinne stellt sich damit als Anwendungssituation dar, die den Kern des aktiven Kompetenzerwerbs kennzeichnet, denn in der Anwendungssituation werden Schülerinnen und Schüler eigenaktiv und führen Denkoperationen und Handlungen aus. Zusammenfassend stellt sich das gezeigte Handeln einer Person als die Beobachtung dar, aus der auf eine bestimmte Kompetenz geschlossen wird. Unabhängig davon, ob die Kompetenz als Disposition verstanden wird und der Performanz vorgelagert ist (Chomsky, 1973, Ziener, 2006) oder ob die Kompetenz als latente Fähigkeit erst durch die Performanz messbar wird und damit Kompetenz erst im Zusammenhang mit konkreten Akten und Fähigkeiten beobachtbar wird, hält Ziener (2006, S. 19) fest: „Es gibt keine Kompetenz ohne Performanz, aber auch kein Handeln ohne Disposition.“. Die Unsicherheiten in der Mehrdeutigkeit der Performanz können minimiert werden, indem die Lernerin oder der Lerner vermehrt mit ähnlichen Problemsituationen konfrontiert wird, die sie oder ihn motivieren, das erworbene Wissen und die vorhandenen Fähigkeiten erfolg-

reich anzuwenden und zu erweitern. Der Aufbau solcher Kompetenzstrukturen und die Möglichkeiten der Abbildung verschiedener Kompetenzfacetten sind Gegenstand des folgenden Abschnitts.

2.1.3 Kompetenzerwerb und Kompetenzaufbau

Die Eigenschaft der Erlernbarkeit von Kompetenzen in Anlehnung an Weinert (2001) impliziert die Möglichkeit, das Kompetenzrepertoire erweitern zu können oder eine neue Kompetenz aufbauen zu können. Den Ausgangspunkt für diese Überlegungen bildet ein Fundament aus Vorwissen, das in seiner Ausprägung nicht in ausreichender Quantität und Qualität vorhanden ist. Weinert (1998) leitet daraus vier Kompetenzbereiche ab, die aufgebaut werden müssen, um den in den Bildungszielen formulierten Anforderungen am Ende der Schulzeit zu genügen. Er konstatiert den Aufbau

1. einer soliden Basis vielfältigen *Wissens*
2. von Strategien zur praktischen *Nutzung und Anwendung* des Wissens (*fachliche Kompetenzen*)
3. und Erwerb von allgemeinen Kompetenzen und Schlüsselqualifikationen sowie das Lernen des Lernens (*überfachliche Kompetenzen*) und
4. eines Systems von kognitiv-motivationalen Handlungs- und Wertorientierungen (*selbstregulative Kompetenzen*),

um aus kognitiven Fähigkeiten gesellschaftlich wertvolle und reflexive Handlungskompetenzen zu entwickeln (Weinert, 1998, S. 115). Zu diesem Vorwissen zählen Leiß und Blum (2006, S. 34) für die Mathematik das Vorhandensein von „...elementaren Fertigkeiten im flüssigen und flexiblen Umgang mit Zahlen und Größen sowie grundlegende Fertigkeiten im Umgehen mit geometrischen Objekten“. Ebenso fußen die Kompetenzen auf sorgfältig und langfristig angelegten inhaltlichen Vorstellungen zu mathematischen Begriffen und Verfahren. Damit entwickeln sich Kompetenzen auf der Basis eines bereits vorhandenen Vorwissens und werden kumulativ aufgebaut und ausdifferenziert. Für den Unterricht bedeutet diese Anschlussfähigkeit von Kompetenzen eine Sicherung des Wissens „nach hinten“ und die Schaffung von Beweglichkeit der neu zu erwerbenden Kenntnisse „nach vorne“ (Lersch, 2010). Didaktisch führt Bruder (2006, S. 136) für einen langfristigen und untereinander vernetzten Kompetenzaufbau in Mathematik zwei Möglichkeiten auf:

- innerhalb eines Schuljahres über verschiedene Unterrichtsthemen [...] hinweg in horizontaler Verknüpfung

- innerhalb einer Leitidee, aber vertikal mit fachlicher Anreicherung angelegt über mehrere Klassenstufen.

Innerhalb eines Schuljahres konkretisiert Bruder (2006) den Kompetenzaufbau für die kognitive Kompetenzfacette durch den schrittweisen Erwerb von neuem Wissen über mathematische Begriffe, Zusammenhänge und Verfahren. Für die Kompetenzfacette des erfolgreichen Lösens von Problemsituationen verweist sie auf ein nachhaltiges Erlernen von geeigneten Vorgehensstrategien zur Unterstützung der prozessbezogenen Kompetenzen. Über mehrere Klassenstufen lassen sich innerhalb einer Leitidee durch das Bewusstmachen und Systematisieren von mathematischen Begriffen und Zusammenhängen Mathematisierungsmuster aufbauen und Vernetzungen knüpfen. Bruder (2006) weist in diesem Zusammenhang auf die Wichtigkeit von kognitiv aktivierenden Aufgabenstellungen, offenen Elementen mit Differenzierungspotential und Reflexionsmöglichkeiten bei der Unterrichtsgestaltung hin. Im Sinne eines langfristigen Kompetenzaufbaus spricht Helmke (2009) von einem „Herunterbrechen der Standards“ auf niedrigere Klassenstufen und auch Lersch (2010) verwendet den Begriff des „Downsizing“. Damit stellt er die distal zu erreichende Kompetenz, wie sie zur Erreichung der Standards gefordert wird, durch den Aufbau von proximalen Teilkompetenzen in Aussicht, d.h. er setzt Zwischenbilanzen auf dem Weg zu den für das Ende der Schulzeit formulierten Standards (Abbildung 2).

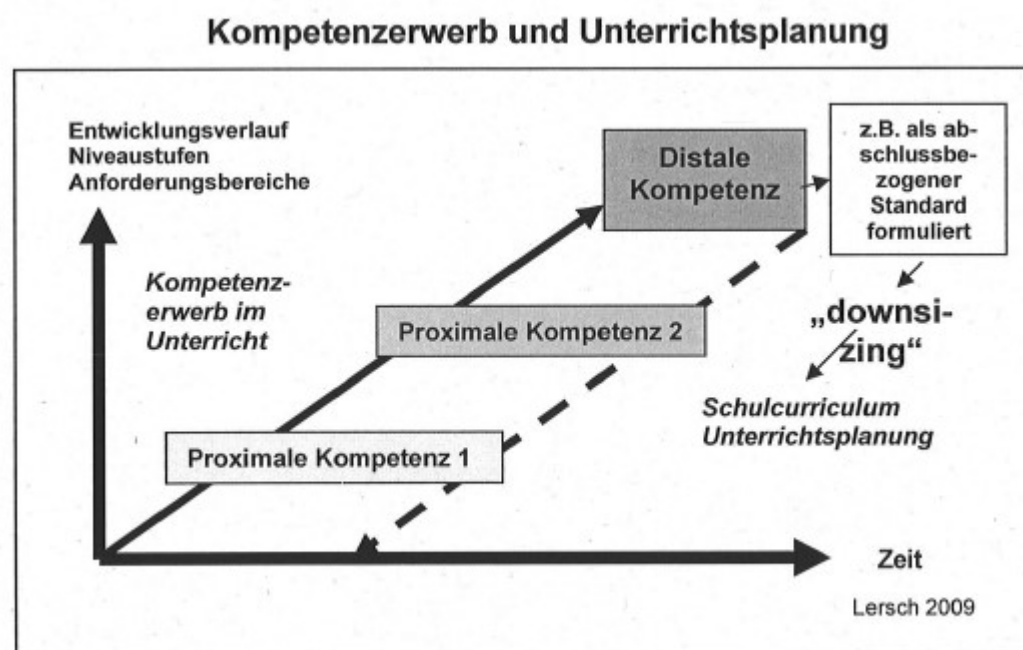


Abbildung 2. Kompetenzerwerb nach Lersch (2010).

der Weg des Kompetenzerwerbs beschrieben werden kann, indem sämtliche Lerngelegenheiten detailliert in eine Reihenfolge gebracht werden. Doch der damit gesetzte Rahmen der Unterrichtsgestaltung führt nicht monokausal zum Kompetenzerwerb. Helmke (2009) führt unter dem Stichwort der „multiplen Bedingtheit des Kompetenzerwerbs“ weitere Faktorbündel auf, die bei der Bilanzierung des Kompetenzaufbaus beachtet werden sollten. Hierunter zählt er sowohl das Lernpotenzial des Schülers, also die kognitiven (Vorwissen, Grundintelligenz, Lernstrategien) als auch die motivationalen Lernvoraussetzungen wie Lernmotivation, Selbstvertrauen und Ausdauer. Einen weiteren Bedingungsfaktor sieht er in der Qualität des Unterrichts, die er wiederum der Lehrperson zuschreibt und nicht zuletzt bezieht er den quantitativen Aspekt des Unterrichts mit ein. „Die verschiedenen Bedingungsfaktoren lassen sich bis zu einem gewissen Ausmaß wechselseitig kompensieren oder substituieren.“ (Helmke & Weinert, 1997 in Helmke, 2009, S. 46). Drieschner (2009) betont in im Zusammenhang der Kumulativität des Lernprozesses die Schaffung didaktischer Strukturierungen, die von einfachen zu komplexen Lernzielen voranschreiten. Aus der Perspektive der Schülerinnen und Schüler verweist er bei der Lernentwicklung auf die starke Abhängigkeit von persönlichen und sozialen Lernvoraussetzungen. Dabei betont Drieschner (2009) in besonderem Maße die Förderung der Schülergruppe mit Migrationshintergrund, die Prinzipien der im Rahmen von pädagogisch-professioneller Lehrerkompetenz zu fokussierenden Individualisierung und Differenzierung sowie die Schaffung von Lerngelegenheiten im Sinne eines adaptiven Unterrichts. Ein Instrument zur Erkennung und Stufung des Könnens sieht Drieschner in dem von Ziener (2006) vorgestellten Verfahren der „Kompetenzexegese“, womit die Aufgabe beschrieben wird, in der kompetenzorientierten Unterrichtsplanung „die Bildungsstandards in erkennbares und gestuftes Können zu übersetzen und somit anschlussfähig an die Heterogenität der Lernentwicklung unter den Kindern einer Lerngruppe zu machen.“ (Drieschner, 2009, S. 76).

Eine weitere Möglichkeit der Abbildung von Kompetenzen und deren Struktur oder Aufbau sehen Klieme et al. (2003) im Potenzial der Kompetenzmodelle. Kompetenzmodelle bieten in erster Linie eine modellhafte Abbildung der Fähigkeiten und Fertigkeiten von Schülerinnen und Schülern, die es auszubilden gilt. Es geht um die Darstellung verschiedener Facetten bzw. Dimensionen einer domänenspezifischen Gesamtkompetenz, wobei der Grad der Ausprägung der Kompetenz vernachlässigt wird. Solche Modelle bilden ein Geflecht theoretischer Konstrukte ab, das die Grundlage für die Operationalisierung und

auch für die Messung der Kompetenzen darstellt. Die Autoren stellen in Aussicht, dass ein Kompetenzmodell Aussagen über die Entwicklung des Zusammenwirkens der verschiedenen Komponenten einer allgemeinen Kompetenz machen könnte, was jedoch mit erheblichen wissenschaftlichen Schwierigkeiten verbunden wäre. Köller (2008) schlussfolgert für die Bildungsstandards, dass nur eine präzise Abbildung des Konstrukts mit allen Facetten und Antezendenzen es ermöglicht, ein tragfähiges Messmodell zu spezifizieren, auf dessen Basis dann Items zur Messung der Kompetenzen entwickelt werden können. Klieme und Leutner (2006) unterteilen die theoretischen Kompetenzmodelle in Kompetenzstrukturmodelle einerseits, in denen es vorrangig um die Definition der Struktur von Kompetenzen und deren Dimensionalität geht und andererseits in Kompetenzniveaus, in denen die Ausprägung einer Kompetenz durch die Ausgestaltung von situativen Anforderungen Rechnung getragen wird. Zur Dimensionalität von Kompetenzen führen Hartig und Jude (2007) aus, dass die bisher in den groß angelegten Vergleichsstudien berechneten eindimensionalen Modelle als Folge auch eine einfache Ergebnisinterpretation nach sich ziehen und von einem einfachen Kompetenzmodell ausgehen. Für eine differenzierte Erfassung von Kompetenzstrukturen oder Kompetenzfacetten kommen mehrdimensionale Modelle zum Einsatz, die es ermöglichen, Teilkompetenzen als separate Dimensionen zu modellieren. Die Abbildung verschiedener Kompetenzniveaus zielt in einer Beschreibung darüber, „...,welche konkreten situativen Anforderungen Personen bei welcher Ausprägung einer Kompetenz bewältigen können.“ (Schott & Ghanbari, 2008, S. 23).

Die Niveaubeschreibungen sind mehrheitlich als Kompetenzstufen bekannt. Für die Bildungsstandards wird der Begriff Kompetenzstufenmodell verwendet, der in Übereinstimmung mit der gängigen Deklaration eine Klassifizierung des Kompetenzkonstruktes vornimmt. Die Fähigkeiten und Fertigkeiten der Lernerinnen und Lerner werden in den Stufenmodellen über ihre Ausprägung differenziert. Ziener (2006) führt für die Bildung von Stufen fünf Vorgehensweisen an und zeigt damit die verschiedenen Konstruktions- und damit zusammenhängend - Interpretationsmöglichkeiten eines Kompetenzstufenmodells. So sieht er *die sachlogisch-graduellen Stufen* als die wahrscheinlichste Form an, in der die Kompetenz mit zunehmendem Fähigkeitsfortschritt zunimmt. Die Stufen der Kompetenzniveaus lassen sich jedoch auch durch die *Verknüpfung von Fähigkeiten* bilden. Hier wird die nächste Kompetenzstufe erst erreicht, wenn bestimmte Verknüpfungen mit an anderer Stelle zu erwerbenden Fähigkeiten erfolgen. Die Kompetenzerweiterung in

die Breite wird auch bei einer *Stufenbildung der Diversifizierung* verfolgt, obwohl es hier gilt, eine Erweiterung des Strategierepertoires zu erzielen und verschiedene Problemlösestrategien zu erlernen. „Die didaktische Frage wird folglich lauten: In welchem Maße kann der oder die Betreffende wählen zwischen unterschiedlichen Lösungsstrategien und eine davon gezielt und sachgerecht einsetzen?“ (Ziener, 2006, S. 38). Bei einem *Experten/Novizen*-Vorgehen wird die Kompetenz der Schülerinnen und Schüler in Relation zu einer vollkommenen Kompetenz gestellt, wodurch sich die Frage stellt, wann ein Schüler oder eine Schülerin ausreichend kompetent ist, um als Experte zu gelten. Schwierig findet Ziener (2006) eine Stufenkonstruktion durch *entwicklungspsychologisch bedingte Stufen*. Er sieht darin die Gefahr, dass eine unterschiedlich ausgeprägte Deutungskompetenz, etwa bei Metaphern in einem Lesetext, zu einer anderen Stufe führen könnte, wenn doch beide Schüler das Optimum an Sinnverständnis im Rahmen ihrer kognitiven Fähigkeiten geboten haben. Im Ergebnis räumt er die Möglichkeiten von sich überschneidenden Modellen ein und stellt heraus, dass der Lernweg ein offener ist und Kompetenzen auf verschiedenen Wegen erworben werden können. Dadurch ordnet er der Kompetenzorientierung im Unterricht den leitenden Stellwert zu: „‘Guter Unterricht‘ wird, auf einen knappen Begriff gebracht, künftig solcher Unterricht sein, der Kinder und Jugendliche im Sinne der Vorgaben ‚befähigt‘ bzw. ihre Befähigung befördert.“ (Ziener, 2006, S. 40).

Aus psychometrischer Sicht schlägt Rost (2004) für eine eindimensionale Messung der Kompetenz drei Arten für eine mögliche Kompetenzstufenbildung auf empirischer Basis vor. Erstens könnte man sich an die Aufgabenschwierigkeiten halten, für die auf der kontinuierlichen Skala gelten würde, dass sich die Schwierigkeiten in den Stufen klumpen, d.h. die Items einer höheren Kompetenzstufe müssten auch schwerer sein als die Items der darunter liegenden Stufe. Für die PISA-Untersuchungen gelang diese Zuordnung nicht vollständig, da eine ganze Reihe von Items in einem anderen Schwierigkeitssegment lagen, als in den a priori angenommenen Kompetenzstufen. Zweitens könnte man anhand der Verteilung der berechneten Personenmesswerte eine mögliche Klumpung identifizieren. „Wenn die Kompetenzstufen so etwas wie homogene Leistungsniveaus darstellen, dann müsste es mehr Personen *innerhalb* jeder Niveaustufe geben als *zwischen* den Stufen.“ (Rost, 2004, S. 665). Die dritte Möglichkeit der Bildung von eindimensionalen Kompetenzstufen sieht der Autor über die Definition der Schülerlösungen. Eine Voraussetzung für diesen Weg sind mehrstufige Aufgaben, die auch eine Teillösung zulassen. Hat ein Schüler oder eine Schülerin nur einen Teil der Aufgabe richtig gelöst, kann diese

Leistung der Kompetenzstufe zugeordnet werden, auf der nur diese Art von Teillösungen definiert ist. Rost (2004) führt mit diesen Vorschlägen psychometrische Möglichkeiten an, die zwar eine willkürliche Setzung von cut-off-Werten (Köller, 2008) vermeiden, letztlich jedoch den qualitativen Unterschieden zwischen den Kompetenzstufen nicht gerecht werden (Rost, 2004). Für die Bildungsstandards wurden schließlich Kompetenzstufenmodelle entwickelt, die in sachlogisch-graduellen Stufen münden, in denen eine kontinuierliche Personenfähigkeit langsam weiterentwickelt wird. Die Stufen sind das Ergebnis von *post hoc* Analysen der gewonnenen Daten und eines Konsens zwischen empirischer Basis und Plausibilitätsargumenten (Köller, 2008). Mit dem Ziel gut interpretierbare Bereiche auf dem Kompetenzkontinuum zu definieren, stellen die Kompetenzstufen der Bildungsstandards ein Zusammenspiel aus psychometrischen Analysen, fachdidaktischen Erwägungen und politischer Verträglichkeit dar (Köller, 2008). Für die Bildungsstandards im Fach Englisch als erster Fremdsprache konnte Tiffin-Richards (2011) zeigen, dass die Experten bei der Setzung von Cut-Scores zur Definition der Kompetenzstufenbreite mehrere Runden des Standard-Settings mit steigender Austausch- und Diskussionsmöglichkeit benötigten, bis eine konsensuale Lösung gefunden wurde. Die aus den Bereichen der Schule, Sprache, Didaktik und Psychometrie zusammengesetzte Teilnehmerrunde veränderte die Setzung der Cut-Scores signifikant von Runde zu Runde und näherte sich dabei einem Konsens an, der an dem fallenden Ausmaß der Standardabweichung abzulesen war. Die Cut-Scores wurden mit steigender Rundenanzahl signifikant höher angesetzt, womit höhere Kompetenzwerte zur Lösung der Aufgaben impliziert werden. Gründe dafür sieht Tiffin-Richards (2011) in der von Runde zu Runde gestiegenen Menge an Informationen, die die Teilnehmer erhielten. So wurde vor der Runde drei eine Verteilung von Schülerleistungswerten über die gesetzten Kompetenzstufen aus Runde zwei bekannt gegeben, woraufhin die Cut-Scores in Runde drei erneut korrigiert wurden. Tiffin-Richards (2011) weist darauf hin, dass die Validität der Kompetenzstufen bei der verwendeten Bookmark-Methode sowohl von der Qualifikation und dem Methodenverständnis der Teilnehmer abhängt als auch von der Qualität der entwickelten Aufgaben, deren Schwierigkeitsgrad auf dem möglichen Kompetenzkontinuum die Grundlage zur Setzung von Cut-Scores bildet. So zeigte die Standard-Setting-Prozedur, dass die Teilnehmer die Cut-Scores für die unterste und oberste Kompetenzstufe so setzten, dass nahezu kein Schüler die Aufgaben im Hörverstehen mit der gesetzten Wahrscheinlichkeit von zwei Dritteln lösen könnte. Der Autor sieht eine Lösung des Problems in der Berichterstattung

von Mittel- oder Medianwerten der einzelnen Panel, deren vorgegebene Items zwar aus einem gemeinsamen Itempool, aber auf unterschiedliche Art und Weise in einem Aufgabenheft zusammengefasst wurden⁸. Er konnte damit zeigen, dass die Wahl der Items in den Testheften der für die Bildungsstandards im Fach Englisch angewandten Bookmark-Methode die Setzung von Kompetenzstufenbreiten beeinflusst: „It is not always possible to employ all available assessment items in standard setting studies, which necessitates the selection of representative item subsamples.“ (Tiffin-Richards, 2011, S. 167).

In der Auseinandersetzung mit dem Kompetenzbegriff zeigt die Vielzahl möglicher Definitionen, wie wichtig eine konkrete Konstruktbeschreibung für die weiterführende Arbeit ist. Die Sicherstellung dessen, was gemessen werden soll, stellt die Grundlage für die späteren Interpretationen dar, weshalb die Definition des Kompetenzbegriffes nach Weinert (2001) im Zusammenhang mit den Bildungsstandards für die weiterführende Arbeit handlungsleitend sein wird. Empirisch gewonnene Personenmesswerte aus einem fachgebundenen Leistungstest sind folglich als die bei den Schülerinnen und Schülern verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten zu verstehen, die notwendig sind, um bestimmte Probleme oder Aufgaben zu lösen.

Für die Bildungsstandards liegen in allen relevanten Domänen theoretische Modelle vor, die mit unterschiedlicher Gewichtung auch empirisch begründet sind⁹. Für das Fach Englisch zeigt Köller (2008) jedoch, dass manche Teilkompetenzen nur schwer zu erfassen sind und sich Teile der Bildungsstandards der Messung ganz entziehen. Ein sowohl theoretisch als auch empirisch gut fundiertes Kompetenzmodell wird für das Fach Mathematik angeboten. Winkelmann et al. (2012) konnten für die Bildungsstandards Mathematik in der Grundschule empirisch nachweisen, dass die mathematische Kompetenz neben einem globalen Faktor auch aus inhaltsbezogenen Dimensionen besteht, so wie sie im Kompetenzmodell für das Fach Mathematik abgebildet sind. Wenngleich die prozessbezogenen Kompetenzen nur unzureichend differenziert werden konnten, wiesen die Autoren nach, dass die Unterteilung von mathematischer Kompetenz in inhalts- und prozessbezogene Kompetenzen dahingehend gerechtfertigt ist, als dass ein fünfdimensionales Modell für die fünf inhaltsbezogenen Kompetenzen eine bessere Datenpassung aufwies

⁸ In der Studie von Tiffin-Richards (2011) erhielt eine Gruppe ein Testheft mit Items, deren Schwierigkeit sich aus den Standardbeschreibungen des Gemeinsamen Europäischen Referenzrahmens für Sprechen ergab. Eine andere Gruppe erhielt ein Testheft mit Items, deren Schwierigkeit empirisch ermittelt wurde. Zur Bookmark-Methode Tiffin-Richards (2011).

⁹ Während sich das Kompetenzmodell der Domäne Deutsch an die PISA-Konzeption anlehnt, wird für die Erste Fremdsprache auf den empirisch wenig belegten „Gemeinsamen Europäischen Referenzrahmen (GER)“ zurückgegriffen.

als ein eindimensionales Modell zur Abbildung einer globalen mathematischen Kompetenz. Die durch die Analysen im Fach Mathematik zur Validierung des Kompetenzmodells vollzogene empirisch abgesicherte Differenzierung der mathematischen Kompetenzen erreichte Pioniercharakter.

2.2 Die Bildungsstandards für das Fach Mathematik in der Sekundarstufe I

Kompetenz- und Kompetenzstufenmodelle stellen für die Anwender von Bildungsstandards zwei grundlegende Instrumente mit strukturierender und diagnostischer Funktion dar. Wenn es den Akteuren durch die Kompetenzmodelle gelungen ist, ein Verständnis für die Zieldimension ihrer pädagogischen Arbeit zu entwickeln, dann bieten ihnen die Stufenmodelle diagnostische Informationen zur weiteren Unterrichtsentwicklung im Sinne der Konstruktion von kognitiv anspruchsvollen Lernarrangements zum weiteren Kompetenzaufbau mit dem Ziel, eine höhere Kompetenzstufe zu erreichen.

Die für die Bildungsstandards entwickelten Kompetenzstufenmodelle entstanden in Folge eines aufwendigen Normierungs- und Standardsetting-Verfahrens, aus dem fächerspezifische Modelle hervorgingen, die für anschließende Untersuchungen gelten und die es erlauben, die Leistungen der Schülerinnen und Schüler in festgelegten Kompetenzstufen zu verorten (z.B. für die landesweiten Lernstandserhebungen und die Ländervergleiche im Sinne eines Bildungsmonitoring gleichermaßen).

2.2.1 Das Kompetenzmodell für das Fach Mathematik

Klieme et al. (2003) ist zu entnehmen, dass das international bekannteste und einflussreichste Rahmenkonzept der National Council of Teachers of Mathematics (NCTM) großen Einfluss auf die Konzeptionsüberlegungen für die Bildungsstandards Mathematik hatte. Die im Jahr 2000 vorgelegten Principles and Standards stellen vielfach beachtete Rahmenbedingungen für den Unterricht mit dem Ziel dar, „eine Leitlinie für die Verbesserung des Mathematikunterrichts im Sinne einer Umorientierung hin zu einem verständnisbasierten und problembezogenen Unterricht zu geben.“ (Klieme et al., 2003, S. 37). Die Principles sind fächerübergreifend und inhaltspezifisch ausformuliert und entsprechen mit ihrer Beschreibung den Methoden und Prinzipien guten Lernens und Lehrens und damit den in Artelt und Riecke-Baulecke (2004) vorgestellten Prozessstandards. Die eigentlichen „Standards“ sind durch ihren fachspezifischen Charakter gute Beispiele für Inhaltsstandards und in zwei Teile gegliedert. Der inhaltliche Teil nennt die gängigen Themenbereiche des Mathematikunterrichts, während der methoden- oder prozessorien-

tierte Teil die wesentlichen Arbeitsfelder aufzählt. Da sowohl die inhaltlichen Themenbereiche wie (1) Zahlen und Operationen, (2) Muster, Funktionen und Algebra, (3) Geometrie und Raumorientierung, (4) Messen und (5) Datenanalyse, Statistik und Wahrscheinlichkeit, als auch die methodenorientierten Aspekte Problemlösen, Argumentieren und Beweisen, Kommunikation, Verbindungen und Darstellungen zusammen unter dem Begriff Standards genannt werden, wird eine Einheit von mathematischen Inhalten und fachbezogenen Arbeitstechniken betont. Klieme et al. (2003, S. 38) sehen in den Standards des NCTM eine Art Kompetenzmodell, dessen fünf Themeninhalte und fünf Prozessaspekte die Gesamtheit mathematischen Denkens und Handelns ausmacht.

Neben der Orientierung an diesen wichtigen Vorarbeiten der NCTM stellen Ehmke, Leiß, Blum und Prenzel (2006) einen weiteren fachspezifischen Aspekt vor, der als Vorlage für die Konzeption der Bildungsstandards im Fach Mathematik diene. Es handelt sich hierbei um das PISA-Rahmenkonzept. Auch PISA beschreibt die Mathematik-Kompetenz aus einer inhaltlichen und einer prozessorientierten Sicht. Diese Differenzierung wurde von den Bildungsstandards aufgegriffen und in die Beschreibung von „mathematischen Leitideen“ und „mathematischen Kompetenzen“ übertragen. Der Einfluss von PISA auf die Bildungsstandards wird vor allem auch bei einem Blick auf die Kompetenzmodelle deutlich. Hier zeigt sich, dass die Bildungsstandards mehrdimensional zu betrachten sind, was durch die Betrachtung der drei Aspekte Prozess-, Inhalts- und Anspruchsdimension zum Ausdruck kommt. PISA unterscheidet „Competencies“ als Prozess-Dimension, „Overarching Ideas“ als Inhalts-Dimension und „Competency Clusters“ als Anspruchs-Dimension. Während die internationale PISA-Rahmenkonzeption sich bei der Operationalisierung durch Aufgaben von den Freudenthal'schen Gedanken (1977) leiten ließ, wonach alles Lehren und Lernen mit der Annahme der Realität als Ausgangspunkt beginnt, lehnt sich die nationale PISA-Konzeption an Winters Grundgedanken (1995) an, wonach die Mathematik als eine deduktiv geordnete Welt eigener Natur verstanden wird. Die internationale Anlehnung an Freudenthal führt zu einer Favorisierung von Vernetzungs- und Modellierungsfähigkeiten, womit die mathematische Aktivität im Vordergrund steht. Als deutsche Erweiterung zu diesem funktionalen Verständnis der Mathematik, fließen Winters grundlegende Arbeiten zum mathematischen Bildungskonzept in die nationale PISA-Konzeption mit ein. Damit wird den curricularen Verhältnissen in Deutschland entsprochen und die funktionale Anwendung mathematischen Wissens um heuristische Fähigkeiten ergänzt, mit denen Kenntnisse zur Problemlösung von Aufgaben ausgebildet

werden, die über die Mathematik hinausgehen. Für die Bildungsstandards im Fach Mathematik bilden die drei Grunderfahrungen nach Winter (2003) die bildungstheoretische Grundlage des Kompetenzmodells. Die Grunderfahrungen (1) Mathematik als Werkzeug, um Erscheinungen der Welt um uns in einer spezifischen Weise wahrzunehmen und zu verstehen, (2) Mathematik als geistige Schöpfung und Welt eigener Art und (3) Mathematik als Hilfsmittel zum Erwerb fachbezogener und fachübergreifender Fähigkeiten, sind in den Standards nicht explizit ausgeführt, dennoch bilden sie eine Klammer um die in den Standards ausgeführten Kann-Erwartungen, die sich nah am mathematischen Arbeiten im Unterricht ansiedeln (Leiß & Blum, 2006).

Für eine produktive Gestaltung des Mathematikunterrichts empfehlen Blum et al. (2006) zusätzlich zu einer Ausrichtung an den mathematischen Kompetenzen, die rahmenden Aspekte um die Bildungsstandards ebenso zu beachten. Hinsichtlich der Formulierung der KMK-Standards im Fach Mathematik führt Reiss (2004, S. 643) jedoch auf, dass man neben den Vorarbeiten und der bildungstheoretischen Verankerung der Standards durch Freudenthal und Winter nicht vergessen sollte, „...dass sowohl die Standards der NCTM als auch die der KMK zu einem gewissen Teil eher implizit als explizit theoriegeleitet sind. Sie können zu einem nicht unerheblichen Anteil als Ergebnisse von *good practice* gesehen werden.“. Dennoch stellen Blum et al. (2006) ein Kompetenzmodell für das Fach Mathematik vor, wie es sich aus den 2003 von der KMK formulierten „Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss“ ergibt.

Das Kompetenzmodell spannt für die Bildungsstandards einen dreidimensionalen Raum auf, dessen Prozess-Dimension die „allgemeinen mathematischen Kompetenzen“ beinhaltet, die Inhalts-Dimension durch „inhaltsbezogene mathematische Kompetenzen geordnet nach Leitideen“ beschrieben wird und die Anspruchs-Dimension durch Anforderungsbereiche klassifiziert (Abbildung 3). Die Dimensionen der Bildungsstandards Mathematik werden im Folgenden in Anlehnung an die Ausführungen von Blum et al. (2006) erläutert.

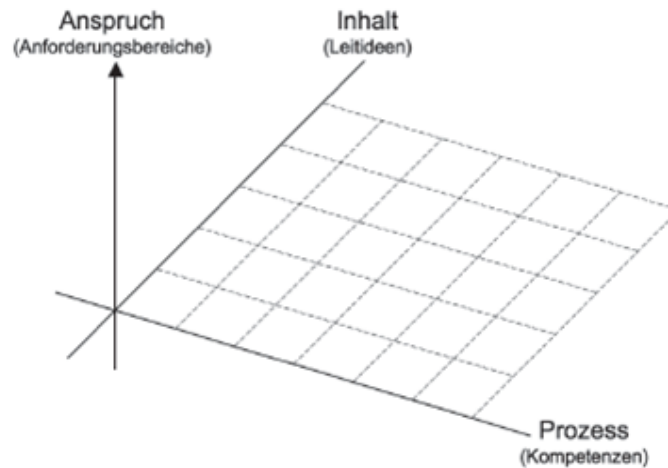


Abbildung 3. Kompetenzmodell für das Fach Mathematik, Sekundarstufe I, Quelle: Blum et al. (2006).

Den Kern der Mathematik-Standards bilden die allgemeinen mathematischen Kompetenzen. Sie bilden die hinreichende Breite des mathematischen Arbeitens ab und treten meist im Verbund mehrerer Kompetenzen auf. Damit wird es nicht möglich, die Kompetenzen scharf voneinander abzugrenzen. Jede der aufgeführten Kompetenzen hat eine aktive und eine passive Komponente. Hier wird deutlich, dass die Ausprägungen nochmals komplexer werden, da es einen Unterschied bildet, selbst Sachverhalte für andere darzulegen oder gegebene Darlegungen aufzunehmen.

Die Anforderungsbereiche sind in drei Stufen unterteilt. Ganz allgemein wird der erste Anforderungsbereich mit „Reproduzieren“ bezeichnet. Der Anforderungsbereich II wird mit „Zusammenhänge herstellen“ überschrieben. Eine Steigerung erfährt der höchste Anspruch, der Anforderungsbereich III, der allgemein „Verallgemeinern und Reflektieren“ genannt wird. Mit Anspruch ist ausdrücklich der Grad an kognitiver Komplexität gemeint, der als Aufgabenmerkmal auch schon für die nationalen PISA-Aufgaben herangezogen wurde.

Blum et al. (2006, S. 35) weisen ausdrücklich darauf hin, dass es „natürlich bei jeder der sechs Kompetenzen auch einen Ausprägungsgrad vor dem Anforderungsbereich I gibt, bei dem die Kompetenz nicht oder jedenfalls nicht in nennenswertem Umfang gefordert wird.“.

Die erste Kompetenz „Mathematisch argumentieren“ beinhaltet die Einsicht, dass bestimmte Begründungsmuster unabhängig vom Inhalt eine gewisse Allgemeingültigkeit haben, da das Spektrum dieser Kompetenz vom Verstehen und kritischen Bewerten verschiedener Formen mathematischer Argumentation bis zum Verbinden mathematischer

Aussagen zu logischen Argumentationsketten reicht. Die Grundlage solcher Argumentationsprozesse sollen fundamentale mathematische Gesetze und Konventionen bilden, womit gleichzeitig ein gewisser Anspruch an Vorwissen geknüpft ist. Neben der typischen Aufgabenstellung mit Begründungs-, Überprüfungs-, Beweis- oder Widerlegungsaufforderung kann eine Aufgabe auch dann Argumentationspotenzial besitzen, wenn bei der Bearbeitung für sich selbst ein Lösungsverfahren oder ein Ergebnis erklärt oder überprüft werden muss. Büchter und Leuders (2009) unterscheiden die Argumentationstätigkeit noch einmal in das außermathematische Argumentieren und das innermathematische Argumentieren. Während die typischen Begründungs- bzw. Beweisketten zum innermathematischen Argumentieren zu zählen sind, findet das außermathematische Argumentieren stets im Rahmen von Modellierungsprozessen statt. So sind die transparenten Darstellungen von Vereinfachungen und Annahmen im Rahmen von Gutachteraufgaben oder Optimierungsproblemen als argumentieren zu verstehen. Die Mindestanforderungen dieser Kompetenz sind jedoch mit der Wiedergabe und Anwendung von Routineverfahren (bekannte Sätze, Verfahren, Herleitungen, usw.), der Wiedergabe rechnerischer Begründungen und dem Argumentieren mit Alltagswissen beschrieben.

Die zweite Kompetenz „Probleme mathematisch lösen“ ist immer dann erforderlich, wenn die Lösungsstruktur einer Aufgabe nicht sichtbar ist und daher eine Strategie für die Bearbeitung entwickelt werden muss. Als Strategieelemente können neben Algorithmen auch heuristische Prinzipien oder Hilfsmittel herangezogen werden. Büchter und Leuders (2009) sehen im Prozess des Problemlösens eine Wahlmöglichkeit aus einer Vielzahl möglicher Verfahren oder der Entwicklung neuer Ansätze. Ebenso ist auch die Modifikation oder Kombination von bekannten Verfahren denkbar. In jedem Fall ist Problemlösen „...ein kreativer Akt, mindestens aber mit Transferleistungen verbunden.“ (Büchter & Leuders, 2009, S. 28). Auf der untersten Stufe kann die Kompetenz beschrieben werden als „Lösen einer einfachen mathematischen Aufgabenstellung durch Identifikation und Auswahl einer nahe liegenden Strategie (z.B. Zeichnen einer einfachen Hilfslinie).“ (Leiß & Blum, 2006, S. 39).

Die Kompetenz „Mathematisch modellieren“ zeichnet sich im Wesentlichen durch die Übersetzungsprozesse aus, die der Schüler beim Bearbeiten einer realitätsbezogenen Fragestellung zu leisten hat. Durch den Einsatz mathematischer Mittel, vor allem mathematischer Modelle, soll der Bearbeiter die Situation verstehen, strukturieren, in die Mathematik übersetzen, eine Lösung finden und danach rückinterpretieren. Büchter und Leuders

(2009) schreiben dem vollständigen Tätigkeitsablauf des Modellierens folgende Schritte zu: die bewusste Auswahl und mathematische Beschreibung des Modells und seiner Annahmen, die Begründung der Auswahl und die mit der Interpretation des Modells verbundene Validierung. Im einfachsten Fall des Anforderungsbereiches I beziehen die Schülerinnen und Schüler vertraute und direkt erkennbare Standardmodelle auf eine Realsituation, wie z.B. den Dreisatz und überführen das Modell direkt in die Mathematik und interpretieren das mathematische Resultat.

„Mathematische Darstellungen verwenden“ ist eine Kompetenz, die beim eigenständigen Erzeugen von Darstellungen mathematischer Gegenstände als auch beim verständigen Umgehen mit gegebenen Repräsentationen benötigt wird. Die Darstellungen und Repräsentationen sind dabei im weitesten Sinne zu verstehen, da auch Formeln, sprachliche Darstellungen, Programme sowie Handlungen und Gesten dazu gehören. Wenn es für den Anforderungsbereich I heißt „Standarddarstellungen von mathematischen Objekten und Situationen anfertigen und nutzen“, dann muss sich der Schüler aktiv mit der Darstellung mathematischer Inhalte auseinandersetzen und die Fähigkeit zum Erstellen einer Darstellung besitzen. Aufgaben, die vom Schüler das Darstellen eines Wahlergebnisses in einem Diagramm verlangen oder das Ablesen von Häufigkeiten aus einem Diagramm sind typische Fragestellungen aus diesem Kompetenzbereich. Aufgaben, die bloße Illustrationen mit motivierender Funktion beinhalten, zählen natürlich nicht dazu.

Die nötigen Werkzeuge zum Bearbeiten vieler Aufgaben werden durch die Kompetenz „Mit symbolischen, formalen und technischen Elementen der Mathematik umgehen“ beschrieben. Hier geht es um die Verwendung der bisher im Unterricht stark vorherrschenden Routinen, die den Gebrauch mathematischer Fakten und Fertigkeiten voraussetzen. Durch Wissen, das direkt aus dem Gedächtnis abgerufen wird und Wissen über mögliche Lösungsverfahren wird diese Kompetenz häufig mit stärker inhaltlich orientierten Kompetenzen kombiniert. Auf der untersten Stufe können die Anforderungen an den Bearbeiter einer solchen Aufgabe wie folgt beschrieben werden: „Verwenden elementarer Lösungsverfahren; direktes Anwenden von Formeln und Symbolen; direktes Nutzen einfacher mathematischer Werkzeuge (z.B. Formelsammlung, Taschenrechner). Typische Aufgaben dieses Anforderungsbereiches fordern die Schülerinnen und Schüler auf, einfache Gleichungen zu lösen oder geometrisch unvollständige Konstruktionen zu vervollständigen.“

Immer wenn es darum geht, einem externen Adressaten Überlegungen, Lösungswege oder Ergebnisse mündlich oder schriftlich zu erläutern, wird die Kompetenz „Mathematisch kommunizieren“ angesprochen. Da die Kompetenz sowohl die Aspekte des Aufnehmens, Verstehens oder Bewertens mathematischer Sachverhalte als auch deren Darlegung in sich vereint, ist die Spanne der kognitiven Anforderungen entsprechend weit zu fassen. Der geringste Anspruch verlangt die Darlegung einfacher mathematischer Sachverhalte durch oder neben der Identifikation und Auswahl von Informationen aus kurzen mathemathikhaltigen Texten, wobei die Ordnung der Informationen im Text weitgehend den Schritten der mathematischen Bearbeitung entspricht. Daher sind Aufgaben, die vom Bearbeiter die Beschreibung einer Rechenoperation verlangen typisch. Beispielhaft sei hier auf Abbildung 4 verwiesen:

Näherungswert für die Kreiszahl π

Bereits in der Antike waren Näherungswerte für die Kreiszahl π bekannt. Damit konnten bei bekanntem Radius der Umfang oder der Flächeninhalt von Kreisen berechnet werden: $U=2\pi r$ bzw. $A=\pi r^2$. Näherungswerte für π erhält man beispielsweise, indem man an konkreten Gegenständen (Rädern, Dosen, ...) den Radius und den zugehörigen Umfang von Kreisen bestimmt. Ermittle experimentell einen Näherungswert für die Kreiszahl π . Dokumentiere dein Vorgehen genau.

Abbildung 4. Die Aufgabe „Näherungswert für die Kreiszahl π “ (Quelle: Blum et al., 2006, S. 218).

Allgemeine Kompetenzen beschreiben auf allen Ebenen die nötigen im Sinne von Weinert definierten Kompetenzen des mathematischen Arbeitens. Diese so genannten Dispositionen zur Bewältigung konkreter Anforderungssituationen eines bestimmten Typs (Klieme et al., 2003, S. 72) werden immer anhand eines konkreten Inhaltsbereiches erworben. In der Konzeption der Bildungsstandards werden die inhaltlichen Bereiche als Leitideen bezeichnet, die versuchen, „die Phänomene zu erfassen, die man sieht, wenn man die Welt mit mathematischen Augen betrachtet.“ (Blum et al., 2006, S. 20). Die Bildungsstandards Mathematik unterscheiden fünf Leitideen, die mit den „Übergreifenden Ideen“ aus der PISA-Konzeption korrespondieren. Jede der fünf inhaltlichen Leitideen Zahl, Messen, Raum und Form, Funktionaler Zusammenhang und Daten und Zufall vereinigt Inhalte verschiedener mathematischer Sachgebiete in sich, die ein mathematisches Curriculum spiralförmig durchziehen (Biehler & Hartung, 2006). Es werden erneut Kompetenzen beschrieben, die inhaltsbezogen sind, jedoch nicht immer eindeutig einer bestimmten Leitidee zuzuordnen sind und Kritikern damit ein schwer wiegendes Argument liefern. So ist für Jablonka und Keitel (2004, S. 138) die Übertragung der mathematischen Leitideen in inhaltsbezogene Kompetenzen keineswegs nachvollziehbar, „...sie erscheint

sogar beliebig“. Wenn die Autoren zu bedenken geben, dass einige der inhaltsbezogenen Kompetenzen auch bestimmten allgemeinen mathematischen Kompetenzen zugeordnet werden könnten, dann folgen sie damit den Konstrukteuren der Bildungsstandards, denn diese betonen ebenso die möglichen Überschneidungen, machen jedoch deutlich, dass die Eindeutigkeit davon abhängt, welcher Aspekt mathematischen Arbeitens im inhaltlichen Zusammenhang betont werden soll (Klieme et al., 2003; Blum et al., 2006).

Die Leitidee „Zahl“ beinhaltet alle Arten von Quantifizierungen, in denen Zahlen verwendet werden, um Situationen zu beschreiben oder zu organisieren. Die nötigen Kompetenzen auf Seiten der Schüler wären demnach die Nutzung sinntragender Vorstellungen von rationalen Zahlen, insbesondere von natürlichen, ganzen und gebrochenen Zahlen entsprechend der Verwendungsnotwendigkeit und die angemessene Darstellung von Zahlen der Situation, z.B. der Zehnerpotenzschreibweise.

Mit der Leitidee „Messen“ werden alle Arten von Größen in verschiedenen Maßeinheiten gemeint, die es zu erfassen und zu verarbeiten gilt. Die Schüler nutzen somit das Grundprinzip des Messens, insbesondere bei der Längen-, Flächen- und Volumenmessung und wählen Einheiten von Größen situationsgerecht aus.

Die Leitidee „Raum und Form“ spricht alle Arten von ebenen oder räumlichen Konfigurationen, Gestalten und Mustern an. Wenn Schüler demnach geometrische Strukturen in der Umwelt erkennen und beschreiben können sowie gedanklich mit Strecken, Flächen und Körpern operieren können, ist die inhaltsbezogene Kompetenz vorhanden.

Alle Arten von relationalen und funktionalen Beziehungen zwischen mathematischen Objekten sind der Leitidee „Funktionaler Zusammenhang“ zuzuordnen. Anwender dieser Kompetenz nutzen Funktionen als Mittel zur Beschreibung quantitativer Zusammenhänge. Sie erkennen und beschreiben diese und stellen sie in sprachlicher, tabellarischer oder grafischer Form dar.

Die fünfte Leitidee „Daten und Zufall“ wurde bislang im Unterricht vernachlässigt. Grund hierfür war die fehlende Betonung in den Lehrplänen und Curricula. Inhaltlich handelt es sich um alle Arten von Phänomenen und Situationen, die statistische Daten beinhalten oder bei denen der Zufall eine Rolle spielt. Kennzeichnend sind hier Tätigkeiten, wie das Auswerten von grafischen Darstellungen und Tabellen aus statistischen Erhebungen, das Beschreiben von Zufallserscheinungen in alltäglichen Situationen, das Bestimmen von Wahrscheinlichkeiten bei einfachen Zufallsexperimenten. Während für den mittleren Schulabschluss auch der Aspekt der Datenanalyse Eingang in die inhaltsbe-

zogenen Kompetenzen fand, wird er in den Standards für die Hauptschule weniger betont (Biehler & Hartung, 2006). Die Abstufung der Inhalte nach Schulformen greifen auch Jablonka und Keitel (2004) in ihrer Kritik wieder auf und bemängeln die unzureichenden Anforderungen an das selbstständige oder kritische mathematische Denken und Argumentieren für die Hauptschule. Auch würde in den Aufgaben das Potential im Sinne des Praxisbezugs z.B. für die Leitidee Funktionaler Zusammenhang nicht genutzt und das unter der Leitidee Raum und Form zusammengefasste Stoffgebiet der Geometrie lehne sich ebenfalls zu eng an mathematische Strukturen an und verspiele damit das Potential aus der außerschulischen Praxis. Inwiefern dem Kritikpunkt der Autoren einer zu kurz gefassten Operationalisierung der Kompetenzen in Testaufgaben zuzustimmen ist, wird in Abschnitt 2.3.3 diskutiert. Leiß und Blum (2006, S. 35) betonen jedoch explizit, dass „es weder sinnvoll noch möglich ist, die Kompetenzen strikt voneinander zu separieren. Bei jeder Kompetenz werden diejenigen Aspekte herausgestellt, die besonders typisch sind. Dennoch sind gewisse Überschneidungen unvermeidbar.“.

Das Kompetenzmodell bildet die Grundlage für die Aufgabenentwicklung zur Überprüfung der Bildungsstandards und auch für die Entwicklung von Lernaufgaben. Es sei jedoch darauf hingewiesen, dass die drei Möglichkeiten zur Klassifikation von Aufgaben (allgemeine Kompetenzen, Leitideen und Anforderungsbereiche) eine theoretische Einordnung in das dreidimensionale Kompetenzmodell erlauben. Spricht man also von erforderlichen Kompetenzen, die für die Bearbeitung einer Aufgabe notwendig sind, so liegt eine kognitive Analyse der Aufgabe auf theoretischer Ebene zu Grunde. Wie die Schüler die Aufgabe letztendlich lösen und welche verschiedenen Lösungswege neben den bisher bekannten noch auftreten, bleibt einer späteren Analyse vorenthalten, so wie Cohors-Fresenborg, Sjuts und Sommer (2004) sie mit den PISA-Aufgaben vorlegten.

2.2.2 Das Kompetenzstufenmodell für das Fach Mathematik

Das auf ca. 500 normierten Items basierende Kompetenzstufenmodell im Fach Mathematik ist eine empirisch entwickelte Abbildung der Standards in Ergänzung zum Kompetenzmodell, das auf eine Kommunikation über den Leistungsstand der Schülerinnen und Schüler abzielt und ebenso die Voraussetzung für eine Definition von Mindest-, Regel und Maximalstandards bildet. Die Festlegungen der Kompetenzerwartungen genügen inhaltlichen, testtheoretischen, fachlichen, curricularen und fachdidaktischen Kriterien, womit konkret eine enge Orientierung an den Bildungsstandards gemeint ist, die zusätzlich aber das gesamte Kompetenzspektrum umfasst. Das Stufenmodell ermöglicht ferner

eine Anbindung an internationale Vorarbeiten, kategorisiert die Kompetenz in fünf gleich breite Stufen mit fachdidaktisch gut interpretierbaren und vertretbaren Grenzen und ergänzt die Leitideenspezifischen Beschreibungen der Kompetenzstufen. Ein solches globales Kompetenzstufenmodell lässt sich auf alle inhaltlichen Kompetenzen anwenden, wobei in der allgemeinen Kompetenzbeschreibung die vorhandenen Kompetenzen betont werden und die damit verbundenen Defizite geschlussfolgert werden.

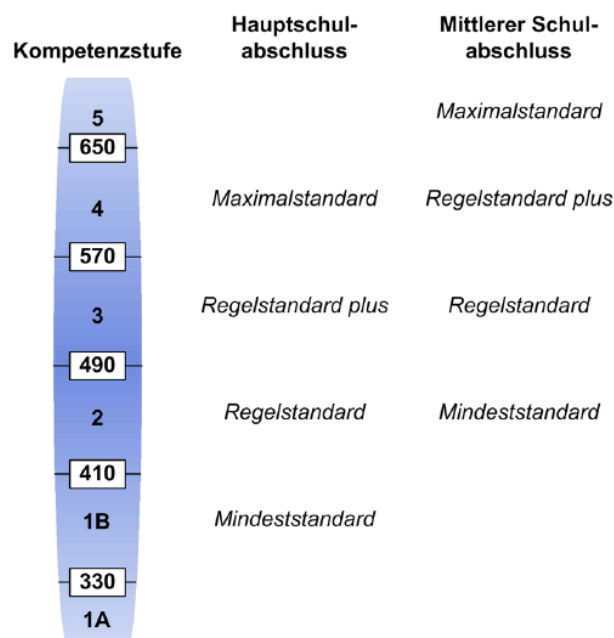


Abbildung 5. Lage von Mindest- und Regelstandards der HSA- und MSA-Population auf dem integrierten Kompetenzstufenmodell Mathematik, Quelle: KMK, 2012.

Die Datenbasis des Kompetenzstufenmodells bilden zum einen ca. 300 Items, die im Rahmen der PISA-Studie 2006 an etwa 10.000 Neuntklässlern normiert wurden. Zusätzlich wurden Testaufgaben für die Jahrgangsstufen 8 und 10 entwickelt, deren Daten in einer Nachnormierung im Umfang von ca. 4.400 Schülerinnen und Schülern den Aufgabenpool ergänzten. Auf Basis dieser repräsentativen Aufgabenlösungen erfolgte die Transformation der Schülerleistungen auf eine Kompetenzskala mit dem Mittelwert 500 und einer Standardabweichung von 100¹⁰ (KMK, 2012). Das Charakteristikum eines solchen Vorgehens besteht darin, dass die aus den Schülerlösungen gewonnene Aufgabenschwierigkeit in direkter Beziehung zu der Fähigkeit der Schülerinnen und Schüler steht und beide Eigenschaften auf derselben Skala abgebildet werden können. Damit ist es

¹⁰ Zum genaueren Vorgehen bei der Skalierung von Kompetenzwerten siehe Abschnitt 5.3.6

möglich, mathematisch-inhaltliche Beschreibungen der Fähigkeiten von Schülerinnen und Schülern aus der Rangfolge der Aufgabenschwierigkeiten und den zu deren Lösung notwendigen Kompetenzen zu generieren. Sinnvolle Kompetenzbeschreibungen entstehen durch die Setzung von Grenzen auf diesem Kontinuum, innerhalb derer die in den Standards beschriebenen Ziele als erreicht gelten. Die Stufensetzung folgt im Rahmen der Arbeit mit den Bildungsstandards dem Prinzip eines kumulativen Kompetenzaufbaus, obwohl eine theoretische Fundierung im Grunde nicht gegeben ist¹¹. Die Stufen sind vielmehr als ein Konsens zu verstehen zwischen verschiedenen Vertretern der Fachdisziplinen und der Bildungspolitik. „Es handelt sich dabei um konsensuelle Festlegungen, welche die Interpretation von Testwerten erleichtern, die aber weit davon entfernt sind, grundlagenwissenschaftlich fundiert zu sein.“ (Köller, 2008, S. 170).

Im Fach Mathematik erfolgte die Festlegung der Kompetenzstufen unter der Leitung des führenden Fachdidaktikers Mathematik im Bereich der Bildungsstandards, Prof. Werner Blum (Universität Kassel) und lehnte sich stark an das Vorgehen bei PISA an. Die Arbeiten im Fach Mathematik stellen durch ihre frühzeitige Umsetzung der Ablaufschritte eines solchen Modellentwicklungsverfahrens die Pionierarbeit dar, an die sich organisatorisch die anderen Fächer anlehnten. Nach anfänglichen fünf Kompetenzstufen ermöglicht es das integrierte Kompetenzstufenmodell, die Aufgabenschwierigkeiten und Schülerkompetenzen in sechs Stufen zu verorten. Die unterste Stufe bleibt dabei nach unten hin offen definiert und die höchste Stufe nach oben hin offen. Die vier mittleren Stufen sind in Intervallen von 80 Kompetenzpunkten gesetzt. Mit festgesetzten Kompetenzstufen durch normierte Aufgaben lassen sich die Schülerleistungen anhand der normierten Aufgabenschwierigkeiten berechnen. Für die Normierungsstudie wurde ein Lernzuwachs von der neunten zur zehnten Jahrgangsstufe von rund 50 Punkten beobachtet. Die Unterschiede der Kompetenzwerte der Neunt- und Zehntklässler führte zu der vorsichtigen Annahme, dass sich ein Lernjahr in etwa 30-50 Kompetenzpunkten ausdrückt (KMK, 2012)¹².

¹¹ Grundsätzlich kann man davon ausgehen, dass die komplexer strukturierten Kompetenzen am oberen Ende ohne die Grundkompetenzen der vorangegangenen Stufen nicht erreicht werden können. Der bisherige Kenntnisstand zur Struktur von Kompetenzen ist hier jedoch noch nicht ausreichend. Köller (2008, S. 172) führt dazu aus: „Die Annahme, es gäbe von Zeit zu Zeit qualitative Sprünge von einer zur nächsten Kompetenzstufe und dies sei entwicklungspsychologisch oder fachdidaktisch plausibel begründbar, steht dabei zumindest teilweise im Gegensatz zu den neueren Erkenntnissen der Lernpsychologie, in der Stufenmodelle keine Rolle spielen.“. Siehe dazu auch die Ausführungen von Helmke und Hosenfeld (2004, S. 63): „Wo es theoretisch überzeugend und empirisch belegbar ist, ist das Stufenkonzept hilfreich. Wo es sich um kontinuierlich verteilte Merkmale handelt, lassen sich auch Abschnitte oder Zonen unterschiedlicher Ausprägung definieren, ohne dass damit die Fiktion von Stufen im engeren Sinne verbunden wäre.“.

¹² Zum Vergleich der Lernzuwächse in anderen Fächern siehe Köller und Baumert (2012).

Eine solche Stufeneinteilung ist ebenso Voraussetzung für die in der Klieme-Expertise propagierte Setzung von Mindest-, Regel- und Maximalstandards, die es den Anwendern erlaubt, eine Diagnose über den Lernstand der Schülerinnen und Schüler herbeizuführen. So kann von einem Erreichen der Mindestanforderungen im Hauptschulbereich erst ab einem Kompetenzwert von 330 Punkten gesprochen werden. Die Mindestanforderungen beim Ziel des Mittleren Schulabschlusses (MSA) liegen bei 410 Punkten (Abbildung 5)¹³. Die Verteilung der Schülerinnen und Schüler der Normierungsstichprobe zeigt, dass 7 % der Hauptschüler (9. Jahrgangsstufe) die Mindestanforderungen im Fach Mathematik verfehlt. Für diese so genannten Risikoschüler besteht ein besonderer Förderbedarf, der sie befähigen soll, in einfachen mathemathikhaltigen schulischen, alltäglichen oder beruflichen Situationen zurechtzukommen. Der Mindeststandard für den Hauptschulabschluss ist mit Kompetenzwerten zwischen 330 und 410 Punkten erreicht und wird durch Aufgaben illustriert, die basale Anforderungen an die Schülerinnen und Schüler stellen. Am Ende der 10. Jahrgangsstufe lassen sich auf dieser Stufe nur vier Prozent der Leistungen der Schülerinnen und Schüler verorten¹⁴. Der Mindeststandard zur Erreichung des Mittleren Schulabschlusses ist auf der Kompetenzstufe 2 ab einem Wert von 410 bis zu einem Kompetenzwert von 490 gesetzt. Für Hauptschüler stellt sie bereits den Regelstandard dar. Diese Kompetenzstufe 2 ist hinsichtlich der allgemeinen mathematischen Kompetenzen gekennzeichnet durch die Durchführung einfacher vertrauter Problemlösestrategien und direkt umsetzbarer Modellierungen, die Verwendung einfacher Darstellungen und die Herstellung von Beziehungen zwischen zwei solchen, die Anwendung wenigschrittiger Standardverfahren und die Fähigkeit, aus Texten bzw. Tabellen einzelne Informationen zu entnehmen. Hinsichtlich der Verteilung der Fähigkeitswerte der ca. 10.000 Schülerinnen und Schüler aus der Normierungsstichprobe zeigt sich, dass 30 Prozent der Neuntklässler und 18 Prozent der Zehntklässler Mathematikkompetenzen auf dieser Stufe nachweisen.

Der Regelstandard für den MSA (und Regelstandard Plus für den Hauptschulabschluss) wird für Schülerinnen und Schüler mit einem Testergebnis auf der Kompetenzstufe 3 erreicht. Aus der Normierungsstudie weiß man, dass sich dort ca. ein Viertel der Neunt-

¹³ Globale Beschreibungen der Kompetenzstufen sind dem von der KMK 2012 verabschiedeten Papier zum „Kompetenzstufenmodell zu den Bildungsstandards für den Hauptschulabschluss und den Mittleren Schulabschluss im Fach Mathematik“ zu entnehmen.

¹⁴ Es bleibt zu beachten, dass bei der Verteilung der Schülerinnen und Schüler der 10. Jahrgangsstufe keine Differenzierung in HSA und MSA vorgenommen wurde, da viele Hauptschüler die Schule nach der 9. Jahrgangsstufe verlassen. Daher wurden keine Hauptschüler in die Stichprobe der Zehntklässler mit aufgenommen (KMK, 2012).

klässler und mehr als ein Drittel der Zehntklässler verorten lassen. Der Regelstandard bezieht sich auf Kompetenzen, die im Durchschnitt von den Schülerinnen und Schülern in einem bestimmten Bildungsabschnitt erreicht werden sollen. Konkret werden mit Kompetenzen auf dieser Stufe folgende allgemeinen mathematischen Kompetenzausprägungen benötigt:

- überschaubare eigene Argumentationen durchführen
- naheliegende Problemlösestrategien anwenden
- wenigschrittige Modellierungen in vertrauten Kontexten vornehmen
- Beziehungen zwischen unterschiedlichen Darstellungen herstellen
- mehrschrittige Standardverfahren oder einschrittige Verfahren mit Variablen anwenden
- überschaubare Lösungswege verständlich darlegen.

Im Hinblick auf die Weiterentwicklung von Unterricht und die Definition eines Leistungsbereichs, der über den Regelstandard hinausgeht, wurde der Regelstandard Plus in Korrespondenz mit der Kompetenzstufe 4 für den MSA festgelegt. Hier können Schülerinnen und Schüler bereits mehrschrittige Argumentationen selbst entwickeln, Problemlösestrategien anwenden, die sie selbst entwickelt haben, mehrstufige Modellierungen vornehmen, eigene Darstellungen erstellen mehrschrittige Operationen mit Variablen ausführen und aus längeren Texten mehrere Informationen entnehmen. Diese durch kreative, komplexe Eigenaktivität gekennzeichnete Kompetenzstufe 4 wird von noch 13 Prozent der Neuntklässler und 30 Prozent der Zehntklässler erreicht.

Die verbleibenden vier Prozent der Schülerinnen und Schüler der neunten Jahrgangsstufe und elf Prozent der zehnten Jahrgangsstufe aus der Normierungsstichprobe verteilen sich auf die Kompetenzstufe 5, mit deren unterster Grenze von 650 Kompetenzpunkten der Maximalstandard beginnt. Die Stufe ist nach oben offen definiert, da dem Kompetenzkontinuum theoretisch in der Komplexität und Anzahl von Denkprozessen keine Grenzen gesetzt sind (*Tabelle 1: Verteilung der Schülerinnen und Schüler der 9. und 10. Jahrgangsstufe auf die Kompetenzstufen, Quelle: KMK, 2012, eigene Darstellung.*). Ausgedrückt in den mathematischen Kompetenzen können Schülerinnen und Schüler auf dieser Stufe komplexe Argumentationen entwickeln und bewerten, anspruchsvolle Problemlösestrategien anwenden und reflektieren, mehrschrittige komplexe Modellierungen vornehmen und beurteilen, komplexe Darstellungen anfertigen bzw. kritisch beurteilen, komplexe innerma-

thematische Verfahren anwenden und kritisch hinterfragen und aus komplexen Texten Sinn entnehmen Informationen gewinnen (KMK, 2012).

Neben der Verabschiedung eines globalen Kompetenzstufenmodells für eine allgemeine mathematische Kompetenz ist es möglich, eine Aufgabentrennung in Leitideen vorzunehmen und damit die mathematische Kompetenz für die fünf inhaltlichen mathematischen Kompetenzen zu differenzieren.

Tabelle 1: *Verteilung der Schülerinnen und Schüler der 9. und 10. Jahrgangsstufe auf die Kompetenzstufen, Quelle: KMK, 2012, eigene Darstellung.*

Kompetenzstufe	Wertebereich	Kriterium	Prozentanteile in der 9. Jahrgangsstufe	Prozentanteile in der 10. Jahrgangsstufe
1	< 410	unter Mindeststandard	28%	< 4 %
2	410 - 490	Mindeststandard	30%	18%
3	490 - 570	Regelstandard	25%	37%
4	570 - 650	Regelstandard Plus	13%	30%
5	> 650	Maximalstandard	4%	11%

Anmerkung: die prozentualen Angaben zur Schülerverteilung wurden der aktuellen Fassung zum Kompetenzstufenmodell von 2012 entnommen

Für das Fach Mathematik in der Sekundarstufe I (HSA und MSA) stehen diese differenzierten Kompetenzstufenmodelle nach Leitideen noch aus. Winkelmann und Robitzsch (2009) konnten für das Fach Mathematik in der Primarstufe nachweisen, dass eine differenzierte Betrachtung mathematischer Kompetenzen nach Inhaltsbereichen als gerechtfertigtes und legitimes Vorgehen zu betrachten ist. Die Autoren verglichen dabei eine eindimensionale Modellierung von 527 normierungstauglichen Items, die 10.328 Schülerinnen und Schülern zur Bearbeitung vorgelegt wurden, mit einer mehrdimensionalen Modellierung, in der jedes Item genau einer inhaltlichen Kompetenz (Between-Item-Dimensionalität) zugeordnet wurde. Im Ergebnis zeigen die Modellgüteindizes einer mehrdimensionalen Modellierung nach inhaltlichen mathematischen Kompetenzen die beste Passung.

Durch die eigene Bedeutung der Kompetenzstufenmodelle in Abgrenzung zu den Kompetenzmodellen werden sie im Rahmen der Kompetenzmessung auf der Grundlage von Bildungsstandards zu dem Kommunikationsmittel, das diagnostische Entscheidungen nach sich zieht und die Gestaltung des Unterrichts in wesentlicher Art beeinflusst. Wenngleich für den Kern der Bildungsstandards, die allgemeinen mathematischen Kompetenzen, aufgrund ihrer nicht trennscharfen Abgrenzung voneinander, keine Kompetenzstufenmodelle

entwickelt werden können, so ist es doch unmittelbar einsichtig, „...dass jede Vorgabe in den Bildungsstandards direkte Konsequenzen für die empirischen Ergebnisse und damit auch für das Kompetenzstufenmodell hat.“ (Reiss & Winkelmann, 2009)¹⁵. Prinzipiell lassen sich die Kompetenzstufenmodelle durch eine erneute Normierungsstudie modifizieren, bisher gilt jedoch das hier beschriebene globale Modell mit den festgelegten Stufenbreiten und inhaltlichen Kompetenzbeschreibungen.

Die Abbildung komplexer Konstrukte wie den Kompetenzen in dimensionsreduzierten Modellen geht gemeinhin mit einem Informationsverlust einher. Trotzdem gilt es, ein handhabbares, kommunizierbares und letztlich doch differenzierendes Modell vorzustellen, das die Realität möglichst nah durch umsetzbare Messbarkeit abbildet. So lassen sich die Aufgaben zur Messung der mathematischen Kompetenzen in einem dreidimensionalen Raum verorten (Kompetenzmodell), während die Berechnung der Testwerte auf eine eindimensionale Abbildung zurückgreift (Kompetenzstufenmodell). Rost (2004, S. 663) beschreibt die Problematik eines am Kriterium der Aufgabenschwierigkeit festgemachten Tests als „...ein Problem der Integration von qualitativer und quantitativer Messung.“. Er macht damit auf die forschungspragmatischen Praktiken aufmerksam, mit denen aktuell die Kompetenzwerte in großen Studien generiert werden und verweist auf die Komplexität der qualitativ unterschiedlichen Denk- und Lernprozesse, die eine Person bei der Lösung von Problemsituationen anwendet. Das Modellieren von qualitativ unterschiedlichen Kompetenzausprägungen auf nur einer Kompetenzdimension ist jedoch durchaus vertretbar, wenn es gelingt qualitativ unterschiedliche Anforderungen in den Testaufgaben zu erfassen, die mit unterschiedlichen Leistungsniveaus korrespondieren (Rost, 2004)¹⁶. In ihrem Modellvergleich zur Dimensionalität mathematischer Kompetenzen konnten Winkelmann und Robitzsch (2009, S. 192-193) für den Primarbereich auf der Grundlage von Modellgüteindizes und Korrelationsmaßen zwischen den Dimensionen zeigen, „...dass der Informationsverlust bei einer eindimensionalen Schätzung mathematischer Kompetenz nicht so erheblich ist, dass diese als Option zu verwerfen wäre – eine Diagnose mathematischer Kompetenz auf Basis eines Gesamtscores oder einer Transformation dieses

¹⁵ Winkelmann und Robitzsch (2009) können zwar zeigen, dass sich auch die prozessbezogenen allgemeinen mathematischen Kompetenzen analytisch voneinander trennen lassen, die Ergebnisse sind aufgrund geringer Itemmengen und einer a priori theoretisch vorgenommenen Klassifizierung der Aufgaben nicht stabil.

¹⁶ Eine Übersicht der Möglichkeiten zur Abbildung und Messung von Kompetenzen geben Klieme und Leutner (2006) in der Beschreibung des Schwerpunktprogramms „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen.“.

Scores etwa in ein globales Kompetenzstufenmodell kann also ebenfalls eine zuverlässige und sinnvolle Methode der Diagnostik darstellen.“.

Für die Arbeiten in Bezug auf die Bildungsstandards geht man aktuell den Weg der ein-dimensionalen Modellierung von Kompetenzwerten und legt somit einen Schwerpunkt der Kompetenzmodellierung auf die Aufgabenentwicklung. Damit wird die Implementation einer neuen Aufgabenkultur zu einer wesentlichen Gelingensbedingung bei der Umsetzung der Bildungsstandards (Abschnitt 2.3.3).

Wenngleich Kompetenzmodelle und Kompetenzstufenmodelle in der Diagnostik von Schülerleistungen eine tragende Rolle spielen, so stellt sich im Hinblick auf den Kompetenzerwerb und –aufbau die Frage, wie ein Unterricht aussehen und gestaltet werden sollte, um eine Weiterentwicklung von Schülerkompetenz zu gewährleisten. Im Folgenden soll es daher auf der Unterrichtsebene darum gehen, welche Merkmale der Unterrichtsgestaltung und –prinzipien zu einem erfolgreichen Kompetenzerwerb und –aufbau führen. Empirische Ergebnisse sollen die theoretischen Ausführungen untermauern und um beobachtetes Unterrichtsgeschehen ergänzen.

2.3 Standard-basierte Unterrichtsgestaltung durch Kompetenzorientierung

Die Ebene des Unterrichts und damit auch die Initiierung von Lernprozessen wird von den Lehrkräften gestaltet. Sie konstruieren Lerngelegenheiten, initiieren Lernprozesse und bestimmen durch ihre Einstellung, Motivation und ihrem Verständnis von Unterricht die Erfolgsfaktoren der Schülerinnen und Schüler bei der Bearbeitung kompetenzorientierter Aufgaben. Dabei versucht die theoriegeleitete Unterrichtsforschung mit dem Konzept des Angebot-Nutzungs-Modell nach Fend (1981), die verschiedenen Traditionen der Schulpädagogik zu vereinigen¹⁷, indem Unterricht als Angebotsstruktur betrachtet wird und die Bereitstellung von Lerngelegenheiten von jedem Lernenden selbst wahrgenommen und genutzt werden muss, um wirksam zu werden. Zur Vorhersage von Lernerfolg im Gefüge von Lernvoraussetzungen, Unterrichtsangebot und Nutzung des Angebots auf Seiten der Schülerinnen und Schüler wird das Modell um spezifische pädagogisch-psychologische Theorien ergänzt. So lassen sich Theorien spezifischer Merkmale der Unterrichtsqualität wie z.B. der Motivation der Lernerinnen und Lerner, dem Professionswissen der Lehrkräfte oder Lernumgebungstheorien und deren Einfluss auf die Leis-

¹⁷ Klieme et al. (2006) referieren als Traditionen der Schulpädagogik (1) das Effizienzmodell eines erfolgreichen, lehrergesteuerten Unterrichts, (2) die Tradition des ganzheitlichen und selbsttätigen Lernens aus der Reformpädagogik und (3) die Tradition des systematischen, fachspezifischen Wissensaufbaus aus der Didaktik.

tung in das Modell integrieren. Klieme et al. (2006, S. 127) führen dahingehend aus, dass Unterrichtsqualität nicht substantiell definiert werden kann, „...sondern umschließt jegliches Merkmal der Lernumgebung, das dazu beitragen könnte, investierte Lernzeit möglichst effizient zu nutzen.“. Im Ansatz von Weinert, Schrader und Helmke (1989, S. 899 in Diedrich et al., 2002) wird Unterrichtsqualität spezifischer verstanden als „jedes stabile Muster von Instruktionsverhalten, das als Ganzes oder durch einzelne Komponenten die substantielle Vorhersage und/oder Erklärung von Schulleistung erlaubt.“. Dass die Schülerleistungen in einem komplexen Beziehungsgefüge durch externe und persönliche Merkmale beeinflusst werden, wird sowohl durch die Large-Scale-Assessments als auch in Experimentalstudien von kleinerem Umfang nachgewiesen. Speziell für den Mathematikunterricht zeigen Köller et al. (2006), dass zwischen schulischen Selbstkonzepten und Interessen ein reziproker Zusammenhang zur Mathematikleistung besteht. Hinsichtlich emotionaler Merkmale wie dem Interesse und der Freude an Mathematik können Pekrun et al. (2006) eine Wechselwirkung mit der Mathematikleistung nachweisen: Emotionen in der 5. Klasse beeinflussen die Notenleistung in der 6. Klasse und die Notenleistung in der 5. Klasse wirkt sich deutlich auf die Emotionswerte in der 6. Klasse aus.

Speziell für den Mathematik- und naturwissenschaftlichen Unterricht spielen Aufgabenstellungen als didaktische Konzeption eine tragende Rolle. Dass in Deutschland Probleme bei der Bearbeitung von Aufgabenstellungen bestehen, wurde durch die TIMSS-Videostudie sichtbar (Klieme et al., 2001), worauf durch eine Fokussierung der Weiterentwicklung der Aufgabenkultur im SINUS-Projekt reagiert wurde (Bund-Länder-Kommission (BLK), 1997). Diese neue Aufgabenkultur ist nach der Einführung der Bildungsstandards, Teil eines als *Kompetenzorientierung*, aber auch *Standard-orientiertes* und *Standard-basiertes Unterrichten* bekannt gewordenen Unterrichtskonzepts. Neben den neuen Aufgaben folgt ein standard-basierter Unterricht dem Prinzip der kognitiven Aktivierung, bietet viele Möglichkeiten des selbstgesteuerten Lernens an und fokussiert auf die in den Standards formulierten Kompetenzen (Helmke, 2009; Lersch, 2010; Oelkers & Reusser, 2008; Ziener, 2006). Ein solcher Unterricht wurde von hinten her gedacht, indem es nicht mehr die Frage zu beantworten gilt, welcher Stoff heute durchgenommen wurde, sondern vielmehr eine Fokussierung auf die aufzubauenden Kompetenzen erfolgt, indem gefragt wird, welche Fertigkeiten und Fähigkeiten heute erworben werden konnten?

Im folgenden Abschnitt werden die angesprochenen Aspekte erneut aufgegriffen und inhaltlich intensiviert. Nach einer allgemeindidaktischen Beschreibung von Kompetenzorientierung wird dargestellt, was konkret die neuen Elemente sind und welche Implikationen damit einhergehen. In einem zweiten Schritt erfolgt die Ausgestaltung der vorangegangenen Betrachtungen konkret für den Mathematikunterricht. Dabei wird die Bedeutung von kompetenzorientierten Aufgaben deutlich gemacht, die in einer weiteren Vertiefung ausgeführt wird.

2.3.1 Unterrichtsverständnis und Unterrichtsskript

Moegling (2010, S. 19) führt als bildungstheoretischen Zugang zur Kompetenzorientierung aus, dass es sich „...vor allem um die Initiierung zur Selbstbildung, zur Förderung des Wunsches sich für die Gestaltung der Zukunft notwendige und auch mitbestimmte und mitverantwortete Kompetenzen anzueignen.“ handelt. Der Autor versteht die Kompetenzorientierung im Allgemeinen als ein von Klafki (1985) hervorgebrachtes Verständnis der Einigkeit von Bildung und Kompetenzorientierung. Klafki (1985) schreibt der Bildung vor allem das Merkmal der Selbstbildung zu, die „...auf der Eigentätigkeit und behutsam angeleiteten Selbststeuerungsprozessen“ beruht (Moegling, 2010, S. 18).

Die schulpädagogische Zielformulierung lässt sich bei der Forderung nach eigentätigen Lernprozessen aus der Tradition der Reformpädagogik herleiten. Inhaltlich sollen die Schülerinnen und Schüler in dieser Tradition neben dem Erwerb von Wissen und Können auch ihre Persönlichkeitsentwicklung vorantreiben. Als didaktisches Ziel eines solchen Unterrichts kann der Aufbau von konzeptuellem Wissen, das problemlösende Denken oder die Nutzung von Wissen in komplexen, authentischen Anforderungssituationen formuliert werden (Klieme et al., 2006). Die Verankerung dieser didaktischen Konzepte folgt einem konstruktivistischen Lehr-Lern-Verständnis, das durch so genannte „diskursive Lernumgebungen“ gekennzeichnet ist und in denen über die Abspeicherung des „trägen Wissens“ hinausgegangen wird (Klieme & Reusser, 2003). Damit zeichnet sich auch das Neue im allgemeinen kompetenzorientierten Unterrichtsskript ab, das konkret in der Annahme besteht, Lernen sei eine eigene, individuelle Konstruktionsleistung. Neue Wissensstrukturen können demnach nur aus der Restrukturierung bereits bestehenden Wissens transportiert werden. Da die Ausgangspunkte des Konstruktionsprozesses bei jedem Schüler verschieden sind, wird Lernen zu einem eigentätig-aktiven Prozess, in dem der Lerner eigenverantwortlich Wissen in mehrheitlich sozial und kulturell geprägten Kontexten erwirbt. Hasselhorn und Gold (2009, S. 233) fassen die Merkmale als wesentliche

Grundbausteine einer moderat konstruktivistischen Lehr-Lern-Philosophie zusammen: „...aktiv, konstruktiv, situiert, selbstregulativ und sozial.“. Die Lernumgebungen in einem solchen Unterrichtsverständnis sind offen, problemorientiert und kognitiv herausfordernd. Im Sinne Aebli's psychologischer Didaktik (1983) soll der Lerner Wissen vom Konkreten zum Abstrakten hin aufbauen, das durch mehrere Anwendungssituationen schließlich über den Transfer für neue Situation anwendbar gemacht wird. Obwohl der Lehrer in Aebli's Theorie eine aktive Position einnimmt, kombiniert er „...Erkenntnisse der verhaltenensorientierten mit der kognitiv-konstruktivistischen Tradition des Lernens.“ (Hasselhorn & Gold, 2009, S. 231).

Im Rahmen der mit der Einführung der Bildungsstandards grundlegenden Ausrichtung des Unterrichts auf Kompetenzorientierung versteht man Unterricht aus einer pragmatischen Perspektive, wie es Reinmann-Rothmeier und Mandl (2001) tun. Konkret wird damit die Koexistenz zwischen Instruktion und Konstruktion in der Instruktionspsychologie beschrieben. Das als gemäßigter Konstruktivismus bezeichnete Lehr-Lernverständnis drückt sich auch in den von Klieme et al. (2006) hergeleiteten Basisdimensionen von gutem Unterricht aus. Die Autoren leiten die Basisdimensionen über das Konzept der kognitiven Aktivität nach Mayer (2004) her und stellen eine Korrespondenz zum Konzept der direkten Instruktion her. „Guter Unterricht, der sowohl die Leistung der Schülerinnen und Schüler im Sinne eines konzeptuellen Verständnisses als auch ihre Motivation fördert, zeichnet sich demnach in drei übergeordneten Merkmalsdimensionen aus:

1. strukturierte, klare und störungspräventive Unterrichtsführung,
2. unterstützendes, schülerorientiertes Sozialklima,
3. kognitive Aktivierung, zu der je nach fachlichem Kontext z.B. [...] herausfordernde, offene Aufgaben in der Mathematik und generell ein diskursiver Umgang mit Fehlern gehören kann (Klieme et al., 2006).

Lersch (2010) macht für die Didaktik eines kompetenzorientierten Unterrichts deutlich, dass es eine Schnittmenge mit dem bisher vorherrschenden inhaltsbezogenen Unterricht gibt. Auch ein kompetenzorientierter Unterricht wird auf der Grundlage eines pragmatischen Unterrichtsverständnisses im Sinne eines gemäßigten Konstruktivismus ohne Inhaltsvermittlung nicht auskommen. In der Praxis soll der stark lehrerzentrierte, gelenkte und eng geführte Unterricht aufgebrochen werden und um kooperative Lernformen, selbstregulierte und kognitiv anregende Lernarrangements erweitert werden.

Ziener (2006) macht deutlich, dass bei der Suche nach einer geeigneten Begriffsdefinition von Kompetenzorientierung ein induktives Vorgehen ratsam ist, wobei aus konkreten Erfahrungen des Lehrens und Lernens ein angemessenes Verständnis abgeleitet wird. Bei einem deduktiven Vorgehen vom Begriff zur Sache würden Lehrerinnen und Lehrer sich herausgefordert fühlen, ihren eigenen bisherigen Unterricht zu verteidigen, obwohl dieser bereits Elemente des neuen Unterrichtsskriptes beinhaltet. Als Instrument zur Strukturierung in der Arbeit mit Bildungsstandards und Kompetenzen rät der Autor zur „Kompetenzexegese“. Ziener (2006) meint damit einen frühen Zeitpunkt in der Unterrichtsplanung, an dem die Lehrkraft die Standards konkretisiert und definiert, was die Schülerinnen und Schüler können müssen, wenn sie die zur Diskussion stehende Kompetenz erworben haben. Das „downsizing“ von distalen Kompetenzen zu proximalen Teilzielen ist in einem kompetenzorientierten Unterricht eine notwendige Technik, um die in den Bildungsstandards nur zu bestimmten Zeitpunkten in einer Bildungskarriere festgeschriebenen can-do-statements auf den Lernprozess der vorgelagerten Lernjahre zu übertragen. Das Erreichen der distalen Kompetenzen (Abschnitt 2.1.3) macht es notwendig, Möglichkeiten für einen Kompetenzerwerb und –aufbau zu schaffen, die mehrheitlich in einer verstärkten Orientierung an konstruktivistischen Unterrichtsentwürfen gesehen wird. Eine Übersetzung in die Didaktik bedeutet, den Unterricht vornehmlich schüler- und handlungsorientiert zu gestalten und damit die Lehrkraft in eine Moderatorenrolle zu verweisen. „Praktisch heißt dies, sich in der Planung und Durchführung des Unterrichts konsequent am Schüler zu orientieren, an seinen Interessen, seinem Vorwissen, seiner Aktivierung etc.“ (Helmke, 2009). Reusser (2006, S. 164) beschreibt die konstruktivistische Lehrerrolle mit einer „Akzentverschiebung von Funktionen und Formen direkter zu solchen eher indirekten Instruktion.“. Dabei macht er deutlich, dass es mit der notwendigen Erweiterung des Methodenrepertoires der Lehrkräfte und einer Öffnung des Unterrichts nicht automatisch zu einem Lernen im konstruktivistischem Sinne kommt, vielmehr wird die Lehrkraft neben einem Moderator je nach Lernumgebung zum Fachexperten, durch individuell abgestimmte, instruktionale Hilfestellung zum Lerngerüst im Sinne eines Scaffolders, zum Dialogpartner oder zum Coach. Reusser (2006) betont im Hinblick auf eine gemäßigte Konstruktivismusauffassung die Wichtigkeit, subjektive Konstruktionsleistungen bei der Lösung einer Problemstellung am Ende wieder einzugrenzen bzw. eine Balance zwischen dem Freisetzen von Subjektivität und dem Reduzieren und Schließen des Denkraums. Im Vergleich zu früheren Unterrichtsgestaltung heißt das, „...von Mono-

kulturen eines auf der Mikroebene lehrergesteuerten Unterrichts wegzukommen und eine Pluralität von Lehr-Lernformen anzustreben.“ (Reusser, 2006, S. 164). In der Evaluation des KMK-Projekts for.mat konnten Asbrand et al. (2012) zeigen, dass es den Lehrkräften weiterhin sehr schwer fällt, den Perspektivwechsel von einer inhaltsbezogenen zu einer kompetenzorientierten Unterrichtsplanung vorzunehmen. Das Festschreiben proximaler Kompetenzen durch schuleigene Curricula oder Arbeitspläne fällt selbst den Schulen mit bereits vorhandener Erfahrung in Schul- und Unterrichtsentwicklungsprozessen schwer. Der noch nicht vollzogene Perspektivwechsel führt in der Auswertung von Gruppendiskussionen dazu, dass Lehrkräfte sich in einem Reflexionsprozess darüber befinden, ob sie die Innovation Bildungsstandards umsetzen sollen.

Abbildung 6 zeigt in einer sehr einfachen Darstellung das Verhältnis von neuem und altem Unterrichtsskript (Lersch, 2010). Mit Verweis auf Ziener (2006) stellt auch Lersch (2010) zwischen den beiden Unterrichtsentwürfen eine Schnittmenge dar, die auf Gemeinsamkeiten schließen lässt. Lersch (2010) vermutet, dass die Schnittmenge noch viel größer ist, da es auch bisher Unterricht gab, der neben der Inhaltsvermittlung darauf angelegt war, die Kompetenzen von Schülerinnen und Schülern zu fördern. Andersherum wird auch deutlich, dass ein kompetenzorientierter Unterricht ebenso Inhalte vermittelt.

Altes vs. neues Unterrichtsskript:



Abbildung 6. Gegenüberstellung von altem und neuem Unterrichtsskript, Quelle: Lersch, 2010.

Die Unterrichtsrealität im Fach Mathematik wurde repräsentativ in der TIMSS-Videostudie beschrieben (Klieme et al., 2001). In einem problemlösenden Mathematikunterricht wurden in deutschen Klassen im Vergleich zu japanischen Klassen eher wenig komplexe Inhalte bearbeitet und in nur geringer Anzahl Denk- und Problemlöseaufgaben gestellt. Die dominierende Unterrichtsform bei der Einführung neuer mathematischer Konzepte ist das Unterrichtsgespräch, das im Vergleich zu Japan sehr eng und lehrergesteuert geführt wird. Dabei zerfällt die anfänglich noch komplexe Aufgabenstellung im Gespräch in sehr kleine Teilaufgaben, die im Ergebnis die Schüler nur noch zu elementaren Tätigkeiten auffordert. „Auf diese Weise wird eine komplexe, offene Problemstellung umgeformt in eine Serie wenig anspruchsvoller, geschlossener Aufgaben.“ (Klieme et al., 2001, S. 45). Für die Schülerinnen und Schüler stagnieren die mathematischen Tätigkeiten auf dem geringen Niveau des Reproduzierens, Assoziierens und dem Anwenden einfacher Operationen, womit eine kognitive Aktivierung auf dem komplexen Niveau des Problemlöseprozesses nicht erfolgen kann. Diese kurzschrittige Erarbeitung spiegelt sich auch in den für Deutschland typischen Unterrichtsskripts wieder, die im Zuge der TIMSS-Videostudie erstellt wurden. Danach folgt der Mathematikunterricht in der Mittelstufe folgendem Muster:

- (1) Die Stunde beginnt mit der Durchsicht und Besprechung der Hausaufgaben.
- (2) Es folgt eine kurze Wiederholungsphase bei zügigem Interaktionstempo.
- (3) Variante 1: Der neue mathematische Stoff wird im fragend-entwickelnden Unterrichtsgespräch, das auf eine einzige Lösung hinführt, relativ kurzschrittig erarbeitet und vom Lehrer an der Tafel dokumentiert.
Variante 2: Wenn das Thema schon in der vorangegangenen Stunde vorbereitet wurde, kann ein Schüler – unterstützt von der Klasse und dem Lehrer – eine Aufgabe an der Tafel entwickeln.
- (4) Anschließend werden in Stillarbeit ähnliche Aufgaben zur Einübung des Verfahrens gelöst.
- (5) Die Stunde schließt mit der Vergabe und Erläuterung der Hausaufgaben (BLK, 1997).

Klieme et al. (2001) schlussfolgern aus einem eng geführten Unterrichtsgespräch, dass die Lehrkraft zwar vage und offene Aufgaben formuliert, jedoch sehr enge Erwartungen für die Antwort hat. Die Schülerinnen und Schüler machen daraufhin häufig rein assoziative Einwürfe, die bei nicht vorhandener Passung in den Gedankengang der Lehrkraft

oftmals unbeantwortet und ohne Feedback bleiben oder verworfen werden. Die Prüfung alternativer Lösungswege erfolgt praktisch nicht. Zu einem positiveren Ergebnis kommen Hugener et al. (2007) bei der Untersuchung von Inszenierungsmustern in den Einführungssequenzen zum Mathematikunterricht in Anlehnung an die Ergebnisse aus der TIMSS-Videostudie. Es zeigte sich, dass die Mehrheit der untersuchten Klassen mit einem problemlösend-entdeckenden Inszenierungsmuster arbeiten, in dem nach einer anfänglichen Problemstellung eine Arbeitsphase der Schülerinnen und Schüler folgt, in der sie selbstständig (paarweise oder in Gruppen) eigene Lösungswege suchen, die anschließend im Plenum vorgestellt und besprochen werden. Der in TIMSS für Deutschland typisch identifizierte problemlösend-entwickelnde Unterricht, in dem die Lösung eines Problems gemeinsam im Klassenverband unter Führung der Lehrkraft in einem fragend-entwickelnden Unterrichtsgespräch erfolgt, wurde von weniger in dieser Studie untersuchten Lehrkräften durchgeführt. Die Forschergruppe konnte jedoch keinen Effekt des als kognitiv-aktivierend eingeschätzten problemlösend-entdeckenden Inszenierungsmusters nachweisen. Damit zeigt sich zwar, dass neben dem für Deutschland postulierten problemlösend-entwickelnden Mathematikunterricht mit geringer kognitiver Aktivierung auch das problemlösend-entdeckende Inszenierungsmuster mit hoch eingeschätzter kognitiver Aktivierung Anwendung findet, ein „guter Unterricht“ jedoch aus mehreren Lernarrangements und Inszenierungsmustern aufgebaut und lern- und leistungsfördernd sein kann (Hugener, Pauli & Reusser, 2007). Pauli und Reusser (2010) untersuchten, ob sich in der Schweiz ein kulturspezifisches Unterrichtsskript rekonstruieren lässt. Sie fanden zunächst, dass sich die Skripte der Mathematiklehrkräfte nach den Unterrichtsphasen unterscheiden. 71,1 Prozent der Lehrkräfte erstellten unterschiedliche Skripte zu einer Einführungsphase oder einer Vertiefungsphase. Für die von den Autoren analysierte Einführungsphase fanden sie zwei unterschiedliche Herangehensweisen. Die Mehrheit der Lehrkräfte verwendet zur Einführung von neuem Stoff ein fragend-entwickelndes Unterrichtsskript. „Die Stunde beginnt mit einer Wiederholung des Vorwissens. Daran anschließend wird der neue Stoff anhand einer Problemstellung fragend-entwickelnd eingeführt und anschließend zum Teil schriftlich festgehalten, beispielsweise mit einem Hefteintrag. Im restlichen Teil der Stunde werden in selbstständiger Schülerarbeit Aufgaben gelöst, vor der Stillarbeit kann noch eine Phase des gemeinsamen Lösens von Aufgaben auftreten, und die Stillarbeit kann für Besprechungen ein- oder mehrmals unterbrochen werden.“ (Pauli & Reusser, 2010, S. 161-162). Bei der Anwendung eines explorierend-

entdeckenden Skripts beginnt die Stunde ebenfalls mit einer Wiederholungsphase und/oder Hausaufgabenkorrektur und/oder Zielangabe, die von der Darstellung einer Problemstellung gefolgt wird. Es folgt eine Phase der Partner- oder Gruppenarbeit, in der die Lernenden versuchen, das Problem zu lösen, eine Gesetzmäßigkeit zu entdecken oder ein Lösungsverfahren zu entwickeln. Anschließend werden die Ergebnisse im Plenum besprochen und diskutiert. Die Lehrkraft hält die wesentlichen Punkte fest, die von den Schülerinnen und Schülern schriftlich dokumentiert werden. Zum Ende werden Aufgaben in selbstständiger Schülerarbeit in Form von Hausaufgaben bearbeitet. Während die Skripte der Einführungsphase nur ca. ein Viertel des gesamten Mathematikunterrichts ausmachen, findet das Skript der Vertiefungsphase, für das kein unterschiedliches Vorgehen gefunden wurde, in über 50 Prozent der gesamten Unterrichtszeit Anwendung. Das Skript der Vertiefungsphase ist durch eine „...mehr oder weniger kurz getaktete Rhythmisierung“ charakterisiert, in der sich Phasen des gemeinsamen und selbstständigen Lösens von Aufgaben abwechseln. Die Besprechung der selbstständig gelösten Aufgaben nimmt dabei ebenfalls einen nennenswerten Anteil ein (Pauli & Reusser, 2010, S. 162). Die Autoren lehnen aufgrund der Ergebnisse die Annahme eines länderspezifischen Unterrichtsskripts, wie sie durch die TIMSS-Videostudie formuliert worden war, für die Schweiz ab. Auch konnten Pauli und Reusser (2010) keinen Zusammenhang zwischen einem explorierend-entdeckendem Unterrichtsskript und einem konstruktivistischem Lehr- und Lernverständnis feststellen. Die Einschränkungen der Studie¹⁸ führen letztendlich jedoch dazu, dass die Autoren eine tendenzielle Übereinstimmung der Ergebnisse der internationalen Videoanalysen mit den Lehrerangaben zu den Unterrichtsskripts insofern erklären, als sich in beiden Studien mehrheitlich das Bild eines eher traditionellen Unterrichts abzeichnet.

Zusammenfassend lässt sich aus einem geänderten Unterrichtsverständnis eine Paradigmenwende des Lernens und der Erkenntnis ausmachen, die den Unterricht in deutschen Schulen durch veränderte Lehr- und Lernformen umgestaltet. Ein Unterrichtsverständnis, das Lernen als einen eigenständigen konstruktiven Aufbau von Wissen versteht, zieht ebenso didaktische Veränderungen nach sich und erfordert die Akzeptanz einer veränderten Lehrerrolle.

¹⁸ Pauli und Reusser (2010) führen an, dass von Lehrkräften, die angeben, ein explorativ-entdeckendes Skript in der Einführungsphase zu verwenden, nur eine Vertiefungsphase aufgezeichnet wurde und die Lehrpersonen im Fall der Übereinstimmung von beschriebener und aufgezeichneter Unterrichtsphase entsprechend ihrer Beschreibung unterrichtet haben könnten (sozial erwünscht).

Aus den theoretischen Überlegungen zur Fokussierung eines veränderten Unterrichtsverständnisses ergibt sich das in Tabelle 2 dargestellte Gesamtbild der Gegenüberstellung von neuem und altem Unterrichtsskript. Die einzelnen Aspekte folgen dabei keiner Reihen- oder Rangfolge und erheben keinesfalls den Anspruch auf Vollständigkeit.

Tabelle 2. Merkmale von neuem und altem Unterrichtsskript, eigene Darstellung

Merkmal	altes Unterrichtsskript	neues Unterrichtsskript
Gemeinsamkeiten	Vermittlung von Inhalten	
	instruktional	
	selbstregulierendes Lernen	
Unterschiede	instruktivistisches Unterrichtsverständnis	gemäßigt konstruktivistisches Unterrichtsverständnis
	Aufbau von deklarativem Wissen	Aufbau von konzeptuellem Wissen
	lehrerzentriert	schülerorientiert
	deduktives Vorgehen	induktives Vorgehen
	hierarchische Lehrerrolle	diskursive Lernumgebungen, Lehrer als Moderator, Fachexperte, Coach, Lerngerüst
	wenig komplexe Teilaufgaben	kognitiv aktivierende Aufgabenstellungen
	abstrakter Aufgabeninhalt	kontextgebundener Aufgabeninhalt

Konkret für den Mathematikunterricht können notwendige Veränderungsaspekte aus den empirischen Ergebnissen der TIMSS-Videostudie abgeleitet werden, um den in den internationalen Vergleichstests erzielten mittleren Leistungsergebnissen deutscher Schülerinnen und Schüler zu begegnen. Für die Unterrichtsgestaltung leiten Klieme et al. (2001, S.76) folgende Punkte her:

- mehr Sicherung von Grundbildung
- mehr inner- und außermathematische Vernetzungen
- weniger Verfahren und Kalküle
- mehr Denkaktivitäten und Eigenkonstruktionen der Schüler
- mehr Reflexionen
- flexiblerer Methodeneinsatz
- deutlichere Trennung von Lernen und Beurteilen.

Wie sich diese Forderungen nach Veränderung durch die Einführung der Bildungsstandards einlösen lassen, wird im folgenden Abschnitt dargestellt, wenn es darum geht, den Unterricht so zu konzipieren, dass die in den Standards formulierten Kompetenzen aufgebaut und erworben werden sollen.

2.3.2 Kompetenzorientierung im Mathematikunterricht

Diese allgemeindidaktische Ausführung konkretisiert sich mit der Notwendigkeit einer fachspezifischen Betrachtung wie sie Köller (2008) für die Implikationen der Bildungsstandards ausführt zu einem Mathematikunterricht, der explizit den Erwerb und die Förderung der in den Standards formulierten Kompetenzen in den Blick nimmt. Damit wird eine Verschiebung der Zielsetzung des Mathematikunterrichts angestrebt, die weniger begriffliches Wissen oder Kalküle zum Gegenstand hat, hin zur Anwendung des Wissens, zur Modellierungs- und Problemlösefähigkeit mit Hilfe der Mathematik (Reiss & Reiss, 2006). Kunter et al. (2005) leiten im Verständnis eines Unterrichts als Gelegenheitsstruktur drei grundlegende Aspekte für aktives und selbstständiges Lernen her, die fächerunabhängig eine befürwortete Orchestrierung diverser didaktischer Strategien und methodischer Grundformen darstellen. Die drei grundlegenden Aspekte werden als (1) Basisdimensionen qualitätvollen Unterrichts¹⁹, (2) kognitive Aktivierung und (3) individuelle und persönliche Unterstützung gekennzeichnet. Jenseits der didaktischen Betrachtung lassen sich für einen kompetenzorientierten Unterricht im Hinblick auf die Unterrichtsqualität verschiedene Dimensionen ausmachen, die den Erfolg der Lernprozesse wahrscheinlich machen. Konkret für den Mathematikunterricht führt Blum (2006b) die drei Komponenten an, die aufgrund ihrer Fachspezifität im Folgenden auch genauer vor dem Hintergrund eines qualitativ hochwertigen Mathematikunterrichts betrachtet werden:

1. eine fachlich gehaltvolle Unterrichtsgestaltung
2. eine kognitive Aktivierung der Lernenden
3. eine effektive und schülerorientierte Unterrichtsführung.

Die drei Komponenten werden im Folgenden weiter ausgeführt und theoretisch unterfüttert sowie um Forschungsergebnisse ergänzt.

2.3.2.1 Fachlich gehaltvolle Unterrichtsgestaltung

Die Komponente der fachlich gehaltvollen Unterrichtsgestaltung beschreibt Blum (2006b) genauer mit dem vielfältigen Angebot von Gelegenheiten, in denen kompetenzbezogene Tätigkeiten im Mittelpunkt stehen. Diese Lerngelegenheiten sollen vielfältige Vernetzungen herstellen, sowohl innerhalb der Mathematik als auch zwischen Mathematik und Realität. Den Schülerinnen und Schülern soll ein Unterrichtsangebot zuteil werden,

¹⁹ Kunter et al. (2005) meinen damit die allgemein lernförderlichen Aspekte des Unterrichts, wie z.B. eine störungspräventive Klassenführung, eine angemessene Geschwindigkeit, Klarheit, Verständlichkeit und Strukturiertheit der Darbietung.

das im Sinne eines verständnisorientierten Lernens in produktiven Übungsformen den mathematischen Kompetenzaufbau ermöglicht. Die Konzeption solcher produktiven Übungsformen ist der Selbstorganisation des Lernens verpflichtet, worin Übungen mit zu entdeckenden Zusammenhängen oder unterschiedlichen Lösungswegen und Begründungsmöglichkeiten fokussiert werden. „Das Ziel auch der produktiven Übungsformen ist gleichermaßen die Ausbildung eines auf Verständnis basierten, schließlich aber automatisierten Beherrschens der jeweiligen Gegenstände.“ (BLK, 1997, o.S.). Ein solcher Unterricht setzt eine kompetente (bezogen auf die Lernprozesse), aspektreiche (bezogen auf die Gegenstände) und authentische (bezogen auf die fachliche Bedeutung der Inhalte) Gestaltung voraus. In der Berichterstattung zur PISA-Studie 2003 führen Baumert et al. (2004) sechs zentrale Grundsätze des verständnisvollen Lernens nach Baumert und Köller (2000) an, die aufgrund der Studienanlage jedoch nicht in der Tiefe operationalisiert werden konnten. In verkürzter Form lässt sich verständnisvolles Lernen mit den folgenden Punkten beschreiben:

Verständnisvolles Lernen

- ist ein aktiver individueller Konstruktionsprozess,
- ist sinnstiftend,
- ist von den individuellen kognitiven Voraussetzungen (vor allem bereichsspezifischem Vorwissen) abhängig,
- erfolgt situiert und kontextuiert,
- wird durch Motivation und metakognitive Prozesse reguliert,
- wird durch kognitive Entlastungsmechanismen unterstützt²⁰.

Pekrun und Zirngibl (2004) führen aus, dass eine entscheidende Voraussetzung für nach- und außerschulisches Lernen die Bereitschaft und Fähigkeit sei, eigenaktiv Lernvorgänge einzuleiten und zu steuern, weshalb das Merkmal der Selbstregulation mit fünf Items in der PISA-Studie 2003 erfasst wurde. Im Ergebnis wird eine Geschlechterdifferenz deutlich, die den Jungen einen höheren Grad an kognitiver Flexibilität und selbstbestimmtem Handeln bescheinigt. Weiterhin zeigt sich, dass das Merkmal Selbstregulation positiv verknüpft ist mit dem Interesse bzw. der Freude an Mathematik. In der Untersuchung des Lernverhaltens wird deutlich, dass deutsche Fünfzehnjährige in einem geringeren Maße Elaborationsstrategien einsetzen. Deutlich häufiger und im Vergleich mit dem internatio-

²⁰ Zur ausführlichen Darstellung Baumert und Köller (2000) und Baumert et al. (2004).

nenal Durchschnitt überdurchschnittlich nutzen die deutschen PISA-Schüler kognitiv rigide, eher oberflächliche Strategien des Wiederholens und Auswendiglernens von Lösungsalgorithmen, was auch in den Leistungen zum Ausdruck kommt, da für den Erwerb von mathematical literacy modellierungsorientierte, kreative und eigenständige Formen des mathematischen Problemlösens und Lernens zentral sind. In den Fortbildungsmaterialien des Hessischen Kultusministeriums (2008, S. 139) wird ausgeführt: „So sollten die Schüler in einem fachlich gehaltvollen Unterricht konsequent zur geistigen Eigenaktivität stimuliert werden und Gelegenheit zum Kompetenzerwerb erhalten, indem sie beispielsweise verstärkt zum Argumentieren, Probleme lösen oder Modellieren angeregt werden.“. In den Untersuchungen zur Bildungsqualität von Schule konnten Otto et al. (2006) erfolgreich ein Trainingsprogramm zur Förderung sachspezifischer und fächerübergreifender (selbstregulativer) Kompetenzen evaluieren. Sie können zeigen, dass die Selbstregulations- und mathematischen Problemlösekompetenzen bei Grundschulern durch ein gezieltes Schüler- und Lehrertraining signifikant gesteigert werden können. Dabei erfuhren die Schülerinnen und Schüler ein siebenwöchiges Training mit je einer Sitzung pro Woche, in denen sie sich mit mathematischen Problemlösestrategien und Selbstregulationsstrategien beschäftigen. Die Lehrkräfte wurden in fünf Sitzungen über Möglichkeiten der Fremdmotivierung, Konzentrations-, Entspannungs- und Bewegungsübungen zur Lernunterstützung sowie dem Modelllernen informiert. Aufbauend auf den Ergebnissen berichten Komorek, Bruder, Collet und Schmitz (2006) über den Inhalt eines Fortbildungskonzeptes, das auf einem Unterrichtskonzept zur Förderung mathematischen Problemlösens und Selbstregulationskompetenzen basiert. Die Kernideen des Unterrichtskonzeptes beschreiben Komorek et al. (2006) im Projekt *Problemlösen und Selbstregulation fördern - Ausbildungsprogramm* (PROSA) durch eine positive Lerneinstellung und Lernmotivation, binnendifferenziertes Arbeiten in leistungsheterogenen Lerngruppen, Sichern von Basiswissen und Können, ein bewusster und reflektierter Umgang mit Mathematikaufgaben, die Vermittlung heuristischer Bildung und die Einbeziehung der Hausaufgaben im Sinne eines selbstregulierten Arbeitens beim Lösen individuell schwieriger Aufgaben. Im Ergebnis zeigt sich eine deutliche Steigerung der Mathematikleistung, insbesondere die leistungsschwächste Gruppe profitierte von dem im Unterricht angewendeten Konzept. Hinsichtlich der Lerneinstellungen und des Lernverhaltens zeichnete sich bei den Schülern jedoch keine positive Entwicklung ab, was durch den Befragungszeitpunkt und die altersspezifischen Besonderheiten erklärt wird.

Für die mathematischen Tätigkeiten des Argumentierens und des Beweisens haben Reiss et al. (2006) in einer Interventionsstudie gezeigt, dass ein aktivierender Unterricht in einem schülerorientierten Lernarrangement die Schülerkompetenzen fördert und auch zu einer Leistungssteigerung, ebenso gerade im unteren Leistungsbereich, führt. Das Prinzip der kognitiven Aktivierung fand sich hier im Grad des selbstregulierten Lernens wieder. Eine erhöhte Selbstregulation und die Verbesserung des Problemlösens war das Ziel einer Studie zur Integration von erfolgreich evaluierten Trainingskonzepten in den Mathematikunterricht (Komorek, Bruder & Schmitz, 2004). Die Forschergruppe macht auf das geringe Interesse der getesteten Schülerinnen und Schüler an der Kompetenz Problemlösen aufmerksam, konnte jedoch nach Anwendung eines gezielten Unterrichtskonzeptes eine Steigerung der Problemlösekompetenz für die leistungsschwachen Schülerinnen und Schüler nachweisen. Eine nur sehr geringe Leistungssteigerung zeigen Hamilton et al. (2003) bei einer Untersuchung des Zusammenhangs zwischen initiierten Schülertätigkeiten durch die Lehrkraft und den Schülerleistungen in Mathematik. Die *reform practices scale* umfasste dabei 15 Items und fragte nach der Häufigkeit der Durchführung bestimmter Schülertätigkeiten, wie z.B. Teilnahme an schülergeführten Diskussionen (participate in student-led discussions), Teilen von Ideen oder Lösen von Problemen in kleinen Gruppen (share ideas or solve problems with each other in small groups) oder Arbeit an der Lösung realer Probleme (work on solving a real-world problem). Bei der Untersuchung der Testergebnisse in Mathematik unterscheiden die Autoren im Antwortformat der Aufgaben. Für alle offenen und acht der zehn geschlossenen Antwortformate zeigte sich, dass eine höhere Leistung mit häufiger berichteten reformorientierten Schülertätigkeiten einhergehen, wobei insgesamt jedoch nur fünf der zehn Koeffizienten signifikant von Null verschieden waren. Die Untersuchung erfolgte im Rahmen der Evaluierung von large-scale reforms in den USA. Die Ergebnisse unterliegen jedoch aufgrund der Besonderheiten der landesweiten Reformbemühungen einer Fülle von Einschränkungen (Abschnitt 3.2).

Der Anspruch einer gehaltvollen Unterrichtsgestaltung in einem kompetenzorientierten Mathematikunterricht wird u.a. durch einen Realitätsbezug eingelöst, der die Tätigkeit des mathematischen Modellierens in den Mittelpunkt stellt (Leuders & Leiß, 2006). Schülerinnen und Schüler sollen die Macht und Grenzen mathematischer Modelle erfahren, bestehende Modelle anwenden und bewerten, aber auch selbstständig Modelle aufstellen und nutzen, die gewonnenen Lösungen interpretieren und die Angemessenheit

der Modellanwendung bewerten. Die Voraussetzungen stellen dabei die Lerngelegenheiten in Form von Aufgaben dar und eine unterstützende Unterrichtsgestaltung von Seiten der Lehrkräfte. Damit einher geht eine für Kritik offene Kommunikationskultur über Lösungsansätze, Fehler und Interpretationen. Der fragend-entwickelnde Unterrichtsstil scheint für diesen Anspruch eine ungeeignete Lösung darzustellen. So wird im Gutachten zu Vorbereitung des SINUS-Programms ausgeführt, dass ein fragend-entwickelnder Unterrichtsstil hinsichtlich der Kommunikationsprozesse ein routiniertes Abspiel von Inszenierungen darstellt, das auf der Mikroebene nicht unbedingt zu einem Verständnis bei den Schülerinnen und Schülern führen muss. Durch die trichterförmig angelegte Kommunikationsstruktur wird die Anknüpfung von Wissen an individuelles Vorwissen meist verengt und die Sinnkonstruktion behindert (BLK, 1997).

Mit dem Zulassen multipler Lösungswege führt Neubrand (2006) eine weitere konkrete Ausführung der gehaltvollen Unterrichtsgestaltung an. „Von Bedeutung ist dies einmal, weil multiple Lösungswege der Mathematik selbst angemessen sind, aber auch weil individuelles Lernen unterstützt wird und der Unterricht auf diese Weise kognitiv reichhaltiger wird.“ (Neubrand, 2006, S. 162). Durch die Darstellung verschiedener Lösungswege erhalten die Schülerinnen und Schüler neue Einsichten in den behandelten Gegenstand und erfahren damit eine „strukturelle Einsicht“, womit die Notwendigkeit des vertikalen Transfers und systematischen Kompetenzaufbaus deutlich wird. Horizontale Vernetzungen sind ebenso möglich, indem dem Lerner durch verschiedene Lösungswege bewusst wird, dass Problemlösungen stets durch Variationen verbessert werden können, „...und dass mathematisch begründete Aussagen über reale Probleme nur so gut sein können wie die zu Grunde gelegten Modelle.“ (Neubrand, 2006, S. 163). Fachlich gehaltvoll werden verschiedene Lösungswege, weil das erlangte Wissen in einer Problemsituation angewendet werden muss und damit aus der ursprünglichen Erwerbssituation gelöst wurde. Mit der gestellten Problemsituation wird dem Lerner bewusst, dass das Wissen nicht an einmalige Situationen gebunden ist und in jeweils anderen Zusammenhängen wieder eingesetzt werden kann. Damit wird an den Lerner ein kognitiver Anspruch gestellt, der Aufschluss auf den Leistungsstand geben kann und das verstehende Lernen in dem Sinne fördert, dass selbst nicht gangbare Lösungswege auf weiterreichende mathematische Inhalte im Voraus verweisen. Inwiefern ein kognitiver Anspruch das Lernen aktiviert und Verstehensprozesse herbeiführt, wird im folgenden Abschnitt erläutert. Pauli und Reusser (2010) befragten in einer Schweizer Studie Mathematiklehrkräfte nach der Häufigkeit des

Einsatzes ihrer Lehrformen im Unterricht. Die Diskussion von Lösungswegen rangierte dabei als eine der zwei Formen, die die Autoren mit einem diskursiven Unterricht (konstruktivistisches Unterrichtsverständnis) verbinden, im häufig eingesetzten Bereich von mindestens ein Mal pro Woche. Am häufigsten setzen die Lehrpersonen jedoch Formen ein, die einem traditionellen Unterricht zuzuordnen sind, wie z.B. Stillarbeit, das Vorlösen von Aufgaben durch eine Schülerin oder einen Schüler, das Abfragen der Lernenden, das Erklären oder Vorlösen von Aufgaben durch die Lehrperson und ein fragend-entwickelndes Lehrgespräch.

Zusammenfassend wird das Qualitätsmerkmal der fachlich gehaltvollen Unterrichtsgestaltung konkretisiert und inhaltlich beschrieben durch die Aspekte

- eines unterstützenden Lernverhaltens seitens der Lehrkraft,
- dem vielfältigen Angebot von Kommunikations-, Problemlöse-, Argumentations- und Modellierungsaufgaben mit Realitätsbezug,
- der möglichst häufigen Initiierung von eigenaktiven Denkprozessen mit der Möglichkeit eines selbstregulierten Lernens und
- eine für multiple Lösungswege, Fehler und Interpretationen offene Kommunikationsstruktur im Unterricht.

2.3.2.2 Kognitive Aktivierung im Mathematikunterricht

Blum (2006b, S. 29) führt zur kognitiven Aktivierung als Qualitätsmerkmal von Mathematikunterricht aus: „...der Unterricht stimuliert geistige Schülertätigkeiten, ermöglicht und ermutigt selbstständiges Lernen und Arbeiten, fördert lernstrategisches Verhalten (*heuristische Aktivitäten*) und fordert stets ein Nachdenken über das eigene Lernen und Arbeiten heraus (*metakognitive Aktivitäten*).“.

Die bisher vorliegenden empirischen Daten geben zweierlei Auskunft. Einerseits konnte durch die TIMS-Videostudie gezeigt werden, dass deutscher Mathematikunterricht vornehmlich in den nicht-gymnasialen Bildungsgängen einem lehrerzentrierten Unterrichtsskript folgt, das durch Abarbeiten von Routinen und der Anhäufung von deklarativem Wissen geprägt ist (Kunter, 2005), andererseits stimmen Lehrkräfte in ihren Selbstausskünften in großer Mehrheit dem normativen Konzept eines selbstständig und diskursiven Lernens in der Mathematik zu (Baumert et al., 2004). Diese Zustimmung basiert auf einer im Rahmen von PISA 2003 durchgeführten Befragung zu den präskriptiven Vorstellungen von Mathematiklehrkräften zum Mathematiklernen. Es konnten hierbei zwei bipolare Faktoren ausgemacht werden, die sich in Kombination mit den erfragten Zieldimensionen

als selbstständiges und diskursives Lernen mit Vertrauen auf mathematische Selbstständigkeit auf der einen Seite darstellen sowie als Gegenpol im rezeptiven Lernen und dem Lernen durch Einschleifen und Vereinfachung. Die weiterhin erfragten Ziele Modellierungsfähigkeit und Selbstregulation werden dem normativen Unterrichtskonzept des selbstständigen und diskursiven Lernens zugesprochen. Diese impliziten Lerntheorien und Ziele der Lehrkräfte verwendeten Baumert et al. (2004) als Skalen für die Untersuchung der erfahrungsbasierten Unterrichtswahrnehmung von Mathematiklehrkräften. Die Modellierung eines Sieben-Faktorenmodells diene als Grundlage für die Rekonstruktion der Unterrichtswahrnehmung von Lehrkräften. Besonderer Wert wurde auf die zuverlässige Abbildung der beiden fachdidaktischen Kerndimensionen einer kognitiv herausfordernden Unterrichtsgestaltung und eines enggeführten Unterrichts gelegt. Im Ergebnis zeigt sich, dass Lehrpersonen, die einen kognitiv aktivierenden Unterricht initiieren, ebenso differenzierend und individualisierend arbeiten, während die Beschreibung eines enggeführten Unterrichts durch häufig direkte Instruktionen und einem erhöhten Interaktionstempo erfolgt. Die damit implizierte Kleinschrittigkeit der kognitiven Bewegungen wird bei einer schulformspezifischen Analyse unter Kontrolle der Schülerfachleistung besonders in den Real- und Hauptschulen beobachtet. Ein kognitiv herausfordernder Unterricht scheint laut Lehrerangaben am ehesten in den Gymnasien realisiert zu werden. Innerhalb der Schulformen bestimmt jedoch das Leistungsniveau der jeweiligen Klasse über die Anwendung differenzierender Maßnahmen und die Komplexität im Weiterlernen. In Klassen, die ein hohes Niveau an mathematischem Vorwissen aufweisen, wird seltener kleinschrittig und enggeführt unterrichtet. Eine Klassifizierung von Lehrkräften in Typen der Unterrichtsinszenierung ergibt vier Profilformen, die sich auf alle Schulformen verteilen und die zwei Extremgruppen erkennen lassen. Den besten Fall beschreibt eine Optimalform, die durch einen effizienten, kognitiv herausfordernden und differenzierten Unterricht mit persönlicher Zuwendung beschrieben wird. Das Spiegelprofil bildet die Problemform eines enggeführten Unterrichts mit sehr niedrigem kognitivem Anspruch und hoher sozialer und persönlicher Distanz. Auf beide Formen fallen jeweils 18 Prozent der befragten Lehrkräfte. Der Hauptteil der Lehrkräfte berichtet von einem Unterricht, der zwischen diesen beiden Polen liegt und versucht, eine Engführung zu vermeiden, aber mit dem Bemühen einer persönlichen Zuwendung und Individualisierung. Die Dimension der kognitiven Aktivierung konnte auch in den Schülerauskünften nachgewiesen werden, wenngleich deren Urteil eher für die Beschreibung von Routinen und die

Einschätzung der Qualität von Interaktionen im Unterricht herangezogen wird (Clausen, 2002, Gruehn, 2000). Allerdings beurteilen die Schülerinnen und Schüler die realisierte kognitive Aktivierung bei der Aufgabenstellung, während die Lehrkräfte eher die didaktische Anlage des Unterrichts beschreiben. Als Experten ihres Unterrichts nehmen die Schüler einen kognitiv aktivierenden Unterricht im Zusammenhang mit der Übernahme der pädagogischen Verantwortung durch die Lehrkraft wahr, d.h. eine selbstständigkeitsfördernde Lernumwelt verbinden die Schüler mit der Wertschätzung und Individualisierung ihrer eigenen Person (Baumert et al., 2004). Kunter et al. (2005, S. 504) beschreiben in den weiterführenden Analysen zum Mathematikunterricht der PISA-Schülerinnen und Schüler den Aspekt der kognitiven Aktivierung in Anlehnung an Baumert und Köller (2000) als Lerngelegenheit, die „...Prozesse des verständnisvollen fachlichen Lernens im Unterricht seitens der Lernenden stimulieren.“. Die zentrale Dimension des verständnisvollen Lernens im Mathematikunterricht wird durch den Faktor „kognitiv herausfordernde Unterrichtsgestaltung“ in einem Sieben-Faktorenmodell abgebildet. Der Faktor beinhaltet Skalenwerte zur Auswahl und Implementation von kognitiv herausfordernden Aufgaben und zur Unterstützung kognitiver Selbstständigkeit. Der Faktor korreliert positiv mit dem ebenfalls bestätigten Faktor der Individualisierung und Differenzierung, einer sozialen und persönlichen Orientierung und besonders hoch mit dem Faktor der erweiterten Lernformen. Interessant ist in dieser Untersuchung, dass auch die Schülerinnen und Schüler die kognitive Aktivierung als ein Merkmal wahrnehmen, das sehr hoch mit einer sozialen Orientierung des Unterrichts, dem Ethos der Verantwortung²¹ und einem angemessenen Interaktions- und Durchnahmetempo korreliert.

Die Befunde zur Unterrichtswahrnehmung im Rahmen von PISA lassen aufgrund der Tatsache mittlerer Schülerleistungen jedoch noch enorme Verbesserungsmöglichkeiten in Richtung Optimalform offen. Insgesamt bleibt die Kritik an der deutschen Unterrichtsrealität eines wenig kognitiv herausfordernden Mathematikunterrichts bestehen, der inhaltlich durch Kalkülhaftigkeit und in der Inszenierung durch Engführung gekennzeichnet ist. Pauli, Drollinger-Vetter, Hugener und Lipowsky (2008) untersuchten im Rahmen des schweizerisch-deutschen Projekts „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ die Zusammenhänge zwischen der methodischen Vorgehensweise beim Begriffsaufbau im Mathematikunterricht und drei Dimensionen eines kognitiv akti-

²¹ Den Faktor benennen Kunter et al. (2005) als „Professionellen Ethos der Lehrkraft“. Der Faktor setzt sich aus Skalen zum Respekt und zur Wertschätzung seitens der Lehrkraft sowie aus Individualisierungsmaßnahmen und wahrgenommener Unterstützung beim Lernen zusammen.

vierenden Unterrichts. Auch diese Autoren gehen von einem konstruktivistischen Lernverständnis aus, fanden jedoch keine Effekte des darstellenden, problemlösend-entwickelnden oder problemlösend-entdeckenden Inszenierungsmuster auf den fachlich-kognitiven Lernertrag.

Messner (2009) weist für einen kognitiv aktivierenden Unterricht auf die wichtige Rolle der Aufgabenkultur hin, in der sich die Prinzipien der aktiven Konstruktion und der Problemorientierung des Unterrichts vereinen. Diese neue Aufgabenkultur findet mit dem SINUS-Projekt einen erfolgreichen Träger im Sinne der Implementation eines kompetenzorientierten Unterrichtskonzeptes. Die Problemorientierung wird dabei als der wichtigste Zugang zu einer kognitiv aktivierenden, eigentätigen Denkleistung angesehen. Beim Einsatz von Aufgaben zur Überprüfung der Bildungsstandards wird das Prinzip der kognitiven Aktivierung in der Kombination aus einer Anforderungsniveauekomponente mit den für die Mathematik geltenden allgemeinen mathematischen Kompetenzen umgesetzt, die jede Aufgabe auf einer drei Mal sechsstufigen Skala nach dem Grad ihrer kognitiven Komplexität verortet (Blum, 2006). Im Allgemeinen lassen sich die Anforderungsbereiche aufsteigend nach ihrem Anspruch bzw. der zu erbringenden Denkschritte ordnen. Somit positioniert sich die Aufforderung zu einem einzigen Denkschritt im Bereich des „Reproduzierens“. Die Bearbeitung eines bekannten Sachverhaltes mit der Anforderung, Fähigkeiten und Fertigkeiten zu verknüpfen stellt sich als kognitiv komplexer dar und ist dem Anspruch „Zusammenhänge herstellen“ zuzuordnen. Die kognitiv komplexeste Form der Denkleistungen wird durch Aufgaben aus dem Bereich „Verallgemeinern und Reflektieren“ beschrieben. Diese Aufgaben fragen mathematische Kompetenzen in der Ausprägung ab, u.a. eigene Begründungen und Interpretationen zu formulieren (Blum, 2006). Weiterhin kann jede Aufgabe einer vorrangig nachgefragten allgemeinen mathematischen Kompetenz zugeordnet werden (Abbildung 3). So ist es möglich, jede mathematische Kompetenz zur Bearbeitung einer Aufgabe schwierigkeitsrelevant nach dem Grad ihres kognitiven Anspruches zu differenzieren. Im Sinne eines standardbasierten Unterrichtens stimuliert der Einsatz kognitiv anspruchsvoller Aufgaben den weiteren Kompetenzerwerb hinsichtlich eines selbstgesteuerten, reflexiven, mathematischen Arbeitens. Eine kompetenzorientierte Aufgabenkultur findet den Zugang zur Aktivierung des Schülers über die Lerngelegenheit, in der sie angewendet wird. Mit offenen Antwortformaten und kooperativen Lernformen wird den Schülerinnen und Schülern der

notwendige Kompetenzaufbau ermöglicht, der sie in Problemsituationen die Kompetenzen abrufen lässt.

Einen anderen Aspekt der kognitiven Aktivierung bringen Klieme et al. (2006) durch die Operationalisierung des Konzepts im Projekt „Pythagoras“ in die Diskussion ein. Sie stellen den Umgang mit Hausaufgaben als einen Indikator für kognitiv aktivierenden Unterricht dar. Die bei den Schülerinnen und Schülern eingesetzte Fragebogenskala erfasste neben der Häufigkeit der Vergabe von Hausaufgaben, ob und wie Lehrpersonen auf Fehler bei den Hausaufgaben eingehen und die Wertschätzung von neuen Lösungswegen und Lösungsprozessen. Aufgrund relevanter Befunde aus der Hausaufgabenforschung erfasste die Forschergruppe weiterhin die Bearbeitungszeit der Hausaufgaben, die Häufigkeit der elterlichen Hilfe und die Häufigkeit der Erledigung der Hausaufgaben. Im Ergebnis konnten vorherige Befunde von Trautwein, Köller und Baumert (2001) repliziert und gezeigt werden, dass die Vergabehäufigkeit von Hausaufgaben einen positiven Effekt auf die Leistung hat, der Umfang elterlicher Hilfe jedoch einen negativen Effekt aufweist. Die bloße Kontrolle der Hausaufgaben im Unterricht zeigt keinen Effekt auf die Leistung, andererseits konnte auf Klassenebene ein signifikant positiver Einfluss des prozessorientierten Umgangs mit Hausaufgaben nachgewiesen werden, womit die Autoren konstatieren: „Das Ausmaß kognitiver Aktivierung – exemplarisch gemessen am Umgang mit Hausaufgaben – hat also einen signifikanten Einfluss auf die Leistungsentwicklung im Verlauf des Schuljahres...“ (Klieme et al., 2006, S. 139).

Eine längsschnittliche Untersuchung über den gesamten Zeitraum der Sekundarstufe I zum Kompetenzerwerb und zum Mathematikunterricht liegt mit dem PALMA-Projekt vor (Pekrun et al., 2006). Als ein zentrales Ergebnis der Analysen im Bereich Modellierungs- versus Kalkülkompetenzen der Schülerinnen und Schüler weist die Forschergruppe das starke Übergewicht der Kalkülkompetenzen aus. Bei gleicher kognitiver Komplexität der Items in der Konstruktion der Subskalen zeigen die Leistungsbefunde, dass es sich bei den Defiziten in den Modellierungskompetenzen nicht um ein jahrgangsstufenspezifisches Phänomen handelt. Lediglich in den Hauptschulen näherten sich die Kompetenzwerte der Schülerinnen und Schüler in den Bereichen Kalkülhaftigkeit und Modellieren an. Auch in der Schülerbefragung zeigen die Werte eine Abnahme des so genannten modellierungsorientierten Unterrichts, wogegen die Mittelwerte hinsichtlich der Kalkülorientierung eine deutlich geringere Standardabweichung aufzeigen und insgesamt sehr viel stabiler sind. Pekrun et al. (2006) verknüpfen die Modellierungskompetenzen mit

dem kognitiven Aktivierungsgehalt des Unterrichts. Aus Sicht der Lehrkräfte konnte in einer Analyse des Gesamtdatensatzes über alle Messzeitpunkte die Ergebnisse von Baumert et al. (2004) hinsichtlich des Dimensionenmodells von Unterricht repliziert werden. Durch die verschiedenen Operationalisierungen des Konzeptes der kognitiven Aktivierung wird die Bedeutung dieses Unterrichtsmerkmals erschwert. Einerseits wird in der Art der Unterrichtsführung eine kognitive Aktivierung erkannt, andererseits werden Aufgaben (als Hausaufgaben oder im Unterricht) bzw. die zu ihrer Bearbeitung notwendigen Kompetenzen als Indikator der kognitiven Aktivierung herangezogen. Damit verschwimmen die Ergebnisse verschiedener Studien durch die Benennung des gemessenen Konstruktes. Ein einheitliches Verständnis eines kognitiv aktivierenden Unterrichts scheint es nicht zu geben, weshalb das Merkmal in seiner Komplexität zunimmt.

2.3.2.3 *Effektive und schülerorientierte Unterrichtsführung*

Blum (2006) vereint in diesem Qualitätsmerkmal die scheinbar gegensätzlichen Kriterien der Unterrichtsgestaltung, sich einerseits konsequent am Schüler zu orientieren und andererseits eine effiziente Bearbeitung des Lernangebots herbeizuführen. Kunter (2005) überträgt das aus dem angloamerikanischen Raum stammende Konzept des *effective classroom management* nach Kounin (1976) in den deutschen Sprachraum mit dem Begriff der „Direkten Instruktion“. Im Kern geht es bei einem effektiven Unterrichtsmanagement darum, den Stundenablauf zügig, gradlinig und gut organisiert umzusetzen und Zeitverzögerungen durch Disziplinprobleme zu vermeiden. Schülerinnen und Schülern soll es durch eine starke Lenkung der Lehrkraft ermöglicht werden, dem Unterricht aufmerksam zu folgen. Das die Unterrichtsquantität ansprechende Ziel, den Schülerinnen und Schülern ein möglichst großes Zeitangebot für die Auseinandersetzung mit mathematischen Inhalten anzubieten, folgt einer festen Sequenz von Abläufen, womit es dem Lerner erschwert wird, eigene Lernaktivitäten auszuwählen, da die Lehrkraft den Unterricht durch das Präsentieren des Lerninhalts, Fragenstellen, das Nennen von Beispielen und Einfordern von Übungen dominiert (Helmke, 1999 in Kunter, 2005). Mit Verweis auf die im Abschnitt 2.3.1 dargestellten Ausführungen zum konstruktivistischen Lernverständnis scheint sich hier eine Diskrepanz abzuzeichnen, die das Qualitätskriterium nach Blum (2006b) widersprüchlich erscheinen lässt. Ein von der Lehrkraft instruierter Wissenserwerb steht im Gegensatz zu der eigentätigen autonomen Konstruktion von Wissen im Sinne eines konstruktivistischen Verständnisses. Kunter (2005) führt dahingehend aus, dass sich die Umsetzung einer effektiven Klassenführung auf den organisatorischen Aspekt der Gestaltung

des Klassenkontextes bezieht und keinesfalls die Qualität der vermittelten Inhalte anspricht. Ein lehrergesteuertes Vorgehen im Unterricht kann ebenso konstruktives und problemorientiertes Arbeiten mit komplexen Sachverhalten ermöglichen. Die Konstrukteure der MARKUS-Studie verstanden eine effiziente Klassenführung in dem von Kunter beschriebenen organisatorischen Sinne und erfassten das Merkmal mit fünf Items. Es zeigt sich, dass die Schülerinnen und Schüler ihren Lehrkräften insgesamt eine Effizienz in der Klassenführung bescheinigen. Lediglich im Bildungsgang des Gymnasiums liegen die Angaben knapp unter dem theoretischen Mittelwert. Helmke et al. (2002, S. 348) sehen in einer effizienten Klassenführung die Voraussetzung für anspruchsvollen Unterricht und führen inhaltlich dazu aus: „Einerseits schafft eine effiziente Klassenführung den notwendigen Rahmen und damit fachliche Lerngelegenheiten (opportunity to learn), andererseits erzeugt sie auf Schülerseite den Eindruck der Wichtigkeit der für den Unterricht vorgesehenen Zeit.“. In diesem Verständnis wird den Schülerinnen und Schülern eine Wertschätzung entgegengebracht, die sich in einem Lernklima ausdrückt, in der die Lehrkraft die gesamte Klasse im Blick hat und nicht nur als Stoffvermittler tätig wird. Helmke et al. (2002) beschreiben die ernsthafte Wahrnehmung der Schüler als Persönlichkeit mit dem Begriff der Schülerorientierung, in die Aspekte der Fürsorge, Verantwortlichkeit und Geduld einfließen.

Mit der Vereinbarkeit von direkter Instruktion im Sinne einer effizienten Unterrichtsführung und Unterrichtsqualität gilt es nun, den Aspekt der Schülerorientierung einzubinden, denn Blum (2006) betont hinsichtlich seines Merkmals der effektiven und schülerorientierten Unterrichtsführung, dass eine störungsarme und fehleroffene Lernatmosphäre geschaffen werden sollte, die sich positiv auf die Einstellungen der Schülerinnen und Schüler auswirkt und dadurch die Leistungen verbessert.

Eine Orientierung am Schüler bedeutet im übertragenden Sinne, ihn dort abzuholen, wo er steht. Das heißt, die Interessen, Neigungen, Werthaltungen, das Vorwissen und die Leistungsfähigkeit des Lernalters in einem ersten Schritt zu diagnostizieren und in einem weiteren Schritt im Prozess des Kompetenzaufbaus zu verwerten. Für die Lehrkraft bedeutet das eine individuelle Anpassung des Lerntempos, der Lerntiefe, des Unterstützungsangebots und Motivationsgrades an die individuellen Lernvoraussetzungen der Schülerinnen und Schüler. Der Umgang mit Heterogenität erhebt den Anspruch, den Lernerfolg jedes Schülers zu optimieren, indem individuelle Unterstützungsleistungen durch differenzierte Aufgabenstellungen angeboten werden.

Die vorangegangenen Vertiefungen zeichnen für den Begriff der Kompetenzorientierung ein vielfältiges Bild, das aufgrund der Vielschichtigkeit von Unterricht nur schwerlich verdichtet werden kann. In der konkreten Unterrichtssituation kann es zu unterschiedlichen Bedeutungen kommen, wie Bruder et al. (2008) aufzeigen. So beschreiben die Autoren Kompetenzorientierung in der Phase der Erarbeitung neuer Inhalte, des Entdeckens und Systematisierens mit folgenden didaktischen Anwendungen:

- genetisches, problemorientiertes Vorgehen;
- Anknüpfung an vorhandene Vorstellungen und Aufbau von Grundvorstellungen;
- Zusammenfassungen und Erarbeitung von Wissensspeichern;
- Reflexionen zum Zugewinn an mathematischen Einsichten, heuristischen Strategien und mathematischen Werkzeugen (Bruder et al., 2008, S. 16).

In der Phase des Übens und Wiederholens, der Vertiefung und Anwendung drückt sich ein kompetenzorientiertes Vorgehen in der Sicherung von Basiswissen und intelligenten Kopfübungen aus. Weiterhin sollen verstehensorientierte Aufgaben, die zum produktiven Üben einladen binnendifferenziert eingesetzt werden und heuristische Hilfsmittel, Strategien und Prinzipien explizit gemacht werden. Auch eine sinnvolle Nutzung der Hausaufgaben drückt für Bruder et al. (2008) ein kompetenzorientiertes Vorgehen aus. Unter Bezug auf die bisherigen Ausführungen kann Kompetenzorientierung im Mathematikunterricht als didaktisches Konzept beschrieben werden, das über die Komponenten eines qualitativ hochwertigen Unterrichts die Schülerinnen und Schüler zu einer breiten allgemeinen Bildung führt, mit der sie in der Lage sind, Problemsituationen selbstständig zu lösen. Bruder, Leuders und Büchter (2008, S. 15) konstatieren: „Insbesondere hat der Begriff der „Kompetenzorientierung“ eine viel größere Tragweite als nur die Orientierung an Texten wie Bildungsstandards oder Kerncurricula, er weist deutlich darauf hin, dass schulische Inhalte und fachliches Arbeiten nicht für die nächste Klassenarbeit, sondern für die Bewältigung von Problemsituationen wichtig sind.“. Durch eine fachlich gehaltvolle Unterrichtsgestaltung stehen vielfältige Lerngelegenheiten zum Kompetenzerwerb im Vordergrund, in denen durch eine kognitive Aktivierung der Lernenden die Aneignung von Wissen und Können in konkreten Anwendungs- und Anforderungssituationen unterstützt wird und durch eine fehleroffene, störungsarme Lernatmosphäre stets eigene Denkkonstruktionen ermöglicht. In einer selbstständigkeitsorientierten Behandlung mathematischer Problemsituationen gilt es für die Lehrpersonen, ein angemessenes Verhältnis zwischen instruktionalen, lehrerzentrierten Methoden und konstruktionalen, schülerorientier-

ten Angeboten zu finden. Die zur Lösung einer mathematischen Problemsituation vom Schüler entwickelten Strategien finden in einem kompetenzorientierten Unterricht insofern Berücksichtigung, dass die Denkprozesse ausgetauscht und reflektiert werden. Die Vielfalt eigener Lernwege und die Vernetzung von Wissens-elementen wird durch den geeigneten Einsatz von Aufgaben gefördert, die fachliche Kompetenzen in ihrer Breite ansprechen. Eine Konkretisierung dieser neuen Aufgabenkultur und deren Abgrenzung zu traditionellen Mathematikaufgaben erfolgt im folgenden Abschnitt.

2.3.3 Neue Aufgabenkultur

Aufgaben im Mathematikunterricht sind nach Bruder (2000) zu verstehen als Aufforderung zur Lerntätigkeit. Im Mathematikunterricht repräsentieren sie durch ihre Einbettung nicht nur eine fachlich gehaltvolle Unterrichtsgestaltung, auch der Grad an kognitiver Aktivierung hängt zu einem großen Teil von der im Unterricht gestellten Problemsituation ab. Nach Bruder (2006) fokussiert eine kompetenzorientierte Unterrichtsgestaltung sowohl auf die Verwendung kognitiv aktivierender Aufgaben als auch auf Lernbedingungen eines effektiven Unterrichts. Im Hinblick auf den Einsatz kompetenzorientierter Mathematikaufgaben beschreibt Sjuts (2006, S. 96) es so: „Repräsentiert man Bildungsstandards durch Aufgaben, so geben die Aufgabenbearbeitungen Aufschlüsse darüber, in welchem Maße betroffene Leitideen schon verankert sind, in welchem Maße Kompetenzen bereits entwickelt sind und in welchem Maße Anforderungen eines bestimmten Bereiches bewältigt werden.“. Blum (2006) weist darauf hin, dass eine kompetenzorientierte Aufgabe nicht per se zur Ausformung, Festigung und Weiterentwicklung von Kompetenzen führt, sondern der pädagogische Einsatz und Umgang mit der Gelegenheit, die entsprechenden Tätigkeiten selbst zu vollziehen, mehr noch, über diese Tätigkeiten zu reflektieren, Lösungswege zu begründen, verschiedene Wege zu vergleichen, Ergebnisse kritisch zu diskutieren u.v.m. Reiss und Reiss (2006) charakterisieren die Aufgaben zu den Bildungsstandards durch das oftmalige Fehlen einer klaren Fragestellung. Schülerinnen und Schüler müssen die Umstände bei der Abwägung einer Entscheidung zur Aufgabenlösung selbst abwägen und damit ihr Denken in funktionalen Beziehungen und die Anwendung dieser Fähigkeit in realen Kontexten zeigen. Büchter und Leuders (2005) belegen den Stellenwert der Konstruktion von Aufgaben, die aktiv-entdeckendes Lernen fördern, mit dem großen Interesse am Themenfeld „Weiterentwicklung der Aufgabenkultur“ im Rahmen des SINUS-Projekts. Für die Autoren kommen die fundamentalen Prinzipien des aktiv-entdeckenden Lernens, des stimmigen Bildes von Mathematik und dem Mathema-

tiklernen mit anderen in guten Aufgaben zum Ausdruck, denen sie die Merkmale der Authentizität, Offenheit und des Differenzierungsvermögens zuschreiben. In den SINUS-Projekten erfuhr die neue Aufgabenkultur eine landesweite Bedeutung und wurde von den teilnehmenden Lehrkräften mit großer Mehrheit zur Bearbeitung gewählt, wodurch das Interesse an der neuen Lernkultur zum Ausdruck kam. Im Folgeprogramm SINUS-Transfer (ab 2003) bearbeiteten laut den Erhebungsdaten aus 2004 von 286 teilnehmenden Fachgruppen 47,2 Prozent das Modul zur „Weiterentwicklung der Aufgabenkultur“. Prenzel (2005, S. 266) beschreibt die inhaltliche Arbeit am Modul zur Aufgabenkultur dahingehend, „...die im mathematischen (und naturwissenschaftlichen) Unterricht verwendeten Aufgaben und Problemstellungen im Hinblick auf ihre Anforderungen, Qualität und ihre didaktischen Funktionen (z.B. Erarbeiten, Üben, Transfer) zu überprüfen und zu verbessern.“. Um einer dominierenden Engführung des Unterrichts zu begegnen, wird im Modul ein Schwerpunkt auf die Entwicklung von Aufgabentypen gelegt, die mehrere Vorgehensweisen und unterschiedliche Lösungsmöglichkeiten zulassen. Durch die Variation von Inhalten, Kontexten und Strukturen soll flexibles Wissen generiert werden und die wenig ausgeprägte vertikale Vernetztheit der Unterrichtsthemen verringert werden. Durch die stärkere Integration der systematischen Wiederholung auch länger zurückliegender Themen soll der Erwerb neuer Routinen zurückgehen und vielmehr verhindert werden, dass in Klassenarbeiten nur der jüngste Unterrichtsstoff abgefragt wird (BLK, 1997, Kap. 9). Traditionelle Aufgaben sind geschlossene, zweckorientierte Arbeitsaufträge, die vordergründig auf den Erwerb mathematischer Begriffe oder die Routinisierung von Techniken und Verfahren ausgerichtet sind. Mit Bezug auf Leuders (2001) führt das Hessische Kultusministerium (2008) aus, dass die Aufgabenstellung traditioneller Aufgaben oft sehr kleinschrittig gegliedert sei mit einem für die Schülerinnen und Schüler eindeutig zu erschließenden Lösungsweg, der oftmals sogar in der Aufgabestellung vorgegeben wird. Der Umgang mit Fehlern, Umwegen oder Näherungslösungen ist im Unterricht nicht etabliert und auch die Lösbarkeit einer Aufgabe steht selten zur Debatte.

Kompetenzorientierte und damit auf den Bildungsstandards beruhende Aufgaben sollen mathematisches Denken auf verschiedenen Niveaus fördern, die Schülerinnen und Schüler zu anspruchsvollen geistigen Aktivitäten und Tätigkeiten stimulieren und häufig zu Reflexionen auf der Metaebene anregen. Die zur Kompetenz „Umgang mit symbolischen, technische und formalen Elementen der Mathematik“ (technisches Arbeiten) auffordern-

den Aufgaben sollen im Unterricht nicht wegfallen, sondern reduziert bzw. zielgerichtet verändert werden (Hessisches Kultusministerium, 2008).

Im Unterricht sind Lehrkräfte darauf angewiesen, beobachtbares Verhalten von Schülern in Form von Worten, Symbolen, Zeichnungen oder mündlichen Äußerungen zu bewerten, indem Mutmaßungen über deren Können angestellt werden. Aufgaben zur Diagnose müssen diesen eher unsicheren Schlussfolgerungen Rechnung tragen und sollten daher in der Regel möglichst umfangreiche Schüleräußerungen einfordern (Büchter & Leuders, 2005). Leuders (2006, S. 87) fügt dem hinzu, dass es wenig Sinn macht, Löser und Nichtlöser voneinander zu unterscheiden, sondern es sich als günstig erweist, wenn eine Aufgabe „...auf verschiedenen Niveaus lösbar ist und damit Auskunft über mögliche Leistungsniveaus gibt.“. Zur Verstärkung der Diagnose ist es auch möglich, sich auf bestimmte Teilkompetenzen zu konzentrieren, um die Aussagen zu fokussieren. Eine nach dem Prinzip richtig/falsch aufgebaute Aufgabe kann dieses Diagnosepotenzial nicht erfüllen, daher ist es für qualitative Rückmeldungen besonders wichtig, Aufgaben mit offenem Antwortformat der Form „Beschreibe dein Vorgehen“, „Gib ein Beispiel“ usw. vermehrt in Klassenarbeiten zu stellen. Eine Öffnung des Antwortformats erhöht das Aussagepotenzial einer Aufgabe dahingehend, dass neben der Frage, ob eine Lösung falsch ist, auch zusätzlich erkannt werden kann, wie der Fehler zustande gekommen ist. Die Aufgabe „Regel“ in Abbildung 7 soll beispielhaft für diese Vorgehensweise Pate stehen.

Regel

Ronald hat den Ausdruck $\sqrt{16+9}$ falsch berechnet. Er rechnete $\sqrt{16+9} = 4 + 3 = 7$.

Berechne den Ausdruck richtig.

Erläutere: Wie ist Ronalds Fehler zustande gekommen?

Abbildung 7. Aufgabe "Regel", Quelle: IQB.

Das Diagnosepotenzial dieser Aufgabe liegt neben der Möglichkeit der Feststellung, ob ein Verfahren beherrscht wird auch darin, ob es verstanden wurde. Büchter und Leuders (2005) weisen gerade auf die Wichtigkeit des Verstehens hin, da dadurch die Sicherheit der Durchführung erhöht wird und der Bearbeiter Handlungskompetenz in ungewohnten Situationen erhält, die nicht mehr zum Standardanwendungsfall des Verfahrens gehören. Dazu ist es wichtig, dass die Aufgabe im Hinblick auf die abzufragende Kompetenz sehr gezielt formuliert ist und nicht durch umfangreiche Aufgabenstellungen oder die Einbettung in einen komplexen Kontext überlagert wird. Büchter und Leuders (2005) bezeich-

nen eine solche Aufgabe hinsichtlich der Kompetenzen im Umgang mit Wurzeln als valide. Freilich muss beim Einfordern von Erläuterungen und Begründungen immer auch darauf geachtet werden, dass nicht jede Aufgabe mit einer solchen Aufforderung versehen wird, da sonst die Gefahr von Standardformulierungen auftritt.

Auch Umkehraufgaben eignen sich hervorragend für die verstehensorientierte Diagnose, da hier der üblicherweise eingenommenen Denkrichtung entgegen getreten wird. Neubrand (2006) betont die dadurch provozierten Probier-Strategien und individuellen Lösungsansätze. Eine typische Umkehraufgabe der neunten Klassenstufe mit geringstem Anforderungsniveau beschreibt die Aufgabe „Raute“ (Abbildung 8).

Raute

Von einer Raute (=Rhombus) ist ihr Flächeninhalt bekannt: 120 cm^2 . Die Seitenlängen und die Längen der beiden Diagonalen sind ganzzahlig (gemessen in cm).

Kannst du eine Raute mit diesen Angaben eindeutig konstruieren? Wenn nein, warum nicht?

Abbildung 8. Aufgabe "Raute", Quelle: Neubrand (2006, S. 174).

Die Lösung dieser Aufgabe ist auf mehreren Wegen möglich. Deren Bewertung richtet sich generell nach ihrem kognitiven Anspruch und sollte von der Lehrkraft vorab vorgenommen werden. Nach Neubrand (2006) ist so eine Einschätzung des jeweiligen Leistungsstandes möglich. Die Unterschiedlichkeit im Lösungsverhalten der Schüler gibt der Lehrkraft einen Eindruck über vorhandene kognitive Fähigkeiten, jedoch gibt es keine allgemeine Regel für die Bewertung. Vielmehr muss jede Aufgabe für sich analysiert werden. Auch die in den Bildungsstandards formulierten Anforderungsbereiche der allgemeinen Kompetenzen können dabei nur Anhaltspunkte bilden und ersetzen keine eigens vorgenommene Aufgabenanalyse mit einer a priori Bewertung der möglichen Lösungswege. Natürlich können die Aufgaben neben der Zuordnung zu bestimmten Typen auch in die bereits erwähnte Klassifizierung entsprechend der Dimensionen des Kompetenzmodells eingeordnet werden. Beide vorgestellten Aufgaben sind auf dem Anforderungsbereich I auszumachen, der allgemein als Reproduzieren charakterisiert wird. Die Aufgabe „Raute“ ist inhaltlich der Leitidee Raum und Form zuzuordnen und erfordert zur Erfüllung der Aufgabenstellung die allgemeinen Kompetenzen „Probleme mathematisch lösen“ und „mathematische Darstellungen verwenden“. Bezogen auf den untersten Anforderungsbereich heißt das, dass die Schüler eine einfache mathematische Aufgabenstel-

lung durch Identifikation und Auswahl einer nahe liegenden Strategie lösen und eine Standarddarstellung von einem mathematischen Objekt anfertigen und nutzen (Leiß & Blum, 2006).

Für die Aufgabe „Regel“ sind solche Tätigkeiten nicht nötig. Hier wird neben der allgemeinen Kompetenz des Problemlösens die Kompetenz „mathematisch argumentieren“ angesprochen, die sich auf dem Anforderungsbereich durch die Wiedergabe und Anwendung von Routineargumentationen auszeichnet. Die Wurzelgesetze sind traditionell der Leitidee „Zahl“ zuzuordnen. Sijts (2006) hält diesbezüglich fest, dass die fachgebundene Gestaltung von Aufgaben nach Leitideen, Kompetenzen und Anforderungsbereichen selbstverständlich sei, doch gerade gegenüber einer rein pädagogischen Sicht von Diagnostik der Hinweis auf die mathematisch-kognitive Verwendung von Werkzeugen zur Erschließung der Welt nochmals zu betonen ist. Zu diagnostischen Zwecken sollten die Aufgaben Aufforderungen an die Schüler enthalten, die sich durch „Erläuterungen in einem Begleittext“, „Stellungnahme und Reflexion“, „Analyse von Fehlvorstellungen“ und „Aufnahme von Ideen“ umschreiben lassen, um den jeweils eigenen Denk- und Verstehensprozess nach außen zu kehren und Verborgenes sichtbar zu machen. Daher ist die Verwendung diagnostischer Aufgaben mit den vorgestellten charakteristischen Merkmalen auch in der Verwendung eines Tests mit der Anlage als System- oder Bildungsmonitoring in einer bestimmten Relation unablässig, wenn differenzierte und qualitative Rückmeldungen über den Leistungsstand der Lerngruppe gemeldet werden sollen.

Schülerinnen und Schüler sollen durch den Einsatz von Aufgaben an ihr Vorwissen anknüpfen können und selbstständige Lernprozesse auslösen. Um Lernerinnen und Lerner zum individuellen Entdecken und eigenständigen Erkunden der im Fach Mathematik relevanten Sinnverknüpfungen aufzufordern, schlägt Leuders (2006, S. 88) folgende Kriterien für Aufgaben vor:

- Offenheit für verschiedene, individuelle Lösungswege,
- keine Engführung durch vielfach „atomisierte“ Teilaufgaben,
- Aktivierung zum Denken (nicht nur zum Handeln),
- ermöglichen von vielfältigen Erfahrungen,
- Orientierung an herausfordernden, aber zugänglichen Problemen,
- zulassen von Schülersprache und vorläufigen Begriffen,
- anstoßen von Kommunikations- und Kooperationsprozessen.

Wenngleich eine Aufgabe nicht alle Kriterien in sich vereinen kann, so verspricht die Öffnung von Aufgaben doch eine Hilfe für die Initiierung von Entdeckungsprozessen. Leuders (2006) empfiehlt darüber hinaus, weiterführende Arbeitsschritte nach einer Erarbeitungsphase wie das Zusammentragen von Einzelergebnissen oder systematische Zusammenfassen von Ergebnissen in Teilen von den Schülerinnen und Schülern selbstständig durchführen zu lassen und als Lehrperson nur dort, wo nötig bestehende Bezeichnungen und Begriffe der Mathematik einzubringen. Langfristige Verknüpfungen und flexibles Wissen wird durch Aufgabenstellungen generiert, die verschiedene Kontexte einbinden, durch Übung und Wiederholung jedoch zu automatisierten Abläufen verhilft, womit Freiräume für andere Denkprozesse geschaffen werden.

Für den Einsatz kompetenzorientierter Aufgaben im Mathematikunterricht konnten Asbrand, Heller und Zeitler (2012) im Evaluationsbericht zum KMK-Projekt for.mat zeigen, dass die Lehrkräfte die Aufgaben als illustrierende Beispiele verstanden und anders als im Fach Deutsch, das Mathematik-Kompetenzmodell eine Akzeptanz erfuhr. In Fachkonferenzen mit Unterrichtsentwicklungserfahrung (z.B. SINUS) wurden die Aufgaben als interessant eingestuft und im eigenen Unterricht erprobt, gleichzeitig wurde von diesen Lehrkräften auch eine Einbettung in größere Unterrichtszusammenhänge gesucht. Insgesamt konnten Zeitler et al. (2012) jedoch keinen Perspektivwechsel der Lehrkräfte von der inhaltsbezogenen zur kompetenzorientierten Unterrichtsplanung ausmachen.

Zeitler et al. (2012) belegen anhand von Gruppendiskussionen mit Lehrkräften, dass ein instruktivistisches Unterrichtsverständnis dazu führt, das Potenzial kompetenzorientierter Aufgaben zu vernachlässigen und den Einsatz der Aufgaben punktuell zu planen. Die Unterrichtsplanung im instruktivistischen Skript besteht aus der Suche nach passenden Materialien zur Vermittlung des geplanten Inhalts. Kompetenzorientierte Aufgaben werden nicht als Chance zur Veränderung des Unterrichts wahrgenommen, sondern verharren in der Bedeutung als Unterrichtsmaterial als Abwechslung zum gewöhnlichen Aufgabenformat. Das Unterrichtsgeschehen orientiert sich noch immer an der Instruktion der Lehrkraft, wodurch auch die Praxis der Unterrichtsvorbereitung und –durchführung nicht grundlegend verändert wird.

Die Konstruktion von Aufgaben zur Überprüfung einer Leistung ist generell zu unterscheiden von Lern- oder Diagnoseaufgaben, die vorwiegend auf das Lernpotenzial und den Anregungsgehalt abzielen. So genannte Testaufgaben müssen daher bestimmten Kriterien genügen, um eine möglichst breite Palette der verfügbaren Kompetenzen zu erfassen.

sen und die Leistung zuverlässig und transparent zu bewerten. Während kompetenzorientierte Aufgaben zum Zweck einer Schulevaluation oder insbesondere einer Individualdiagnostik eher als eine Untersuchung der Tiefe der vorhandenen Kompetenzen zu betrachten sind, kann bei einem Test zum Zwecke des Systemmonitorings das Spektrum der erfassten Leistungen drastisch verbreitert werden, da jeder Schüler nur noch einen Teil der Testaufgaben bearbeitet (Klieme et al., 2003). Die Schwierigkeit der Aufgaben wird wie in allen Leistungsvergleichen üblich variiert, kann bei Bedarf und Verwendung so genannter Anker- oder Brückenitems für bestimmte Lerngruppen auch Niveau-angepasst werden, entsprechend ihrem erwartetem Leistungsniveau. Die Verwendung probabilistischer Testmodelle ermöglicht dann trotzdem eine Abbildung der Leistungswerte der Schüler auf einer Skala. Die Forschergruppe um Neubrand arbeitete sechs Aufgabenmerkmale heraus, die gemeinsamen Einfluss auf die Schwierigkeit im Sinne von Lösungshäufigkeiten nehmen: Kontext, Offenheit der Aufgabe, Notwendigkeit einer Begründung, Bearbeitungsumfang, kognitive Komplexität und curriculare Wissensstufe.

Büchter und Leuders (2005) stellen sechs Kriterien vor, die bei der Einschätzung von Aufgaben für die Leistungsbewertung zu berücksichtigen sind. Die *Validität* der Aufgaben zielt darauf ab, genau die Kompetenzen in den Blick zu nehmen, die bewertet werden sollen. Testaufgaben sollen einfach und mit explizit klar gestelltem Bearbeitungsauftrag formuliert werden. Durch die reduzierte Verwendung mathematischer Begriffe soll die *Verständlichkeit* einer Aufgabe weiter erhöht werden. Damit einher geht das dritte Kriterium, die *Erwartungstransparenz*. Für die Schüler soll klar ersichtlich sein, was von ihnen verlangt wird und wann die Bearbeitung die Anforderungen in welchem Maße erfüllt. Dabei muss auch ein Augenmerk auf das passende Verhältnis von Bearbeitungszeit und bewertbaren Schülerlösungen gelegt werden, um das Kriterium der *Ergebnisorientierung* in den Blick zu nehmen. Der Aufgabenkonstrukteur, in aller Regel wird das eine Lehrkraft sein, muss sich bewusst machen, ob die Aufgabe das sichere Beherrschen eines Verfahrens oder das Verständnis eines Verfahrens erfragen möchte. Diese *Verfahrens-* vs. *Verstehensorientierung* hat auch Einfluss auf die Bestimmung der *Schwierigkeit*. Die Faktoren, die eine Aufgabe schwieriger machen, müssen dem Konstrukteur bekannt und vertraut sein. Blum (2006) fügt zu den genannten Kriterien noch die verlässliche Korrigierbarkeit hinzu, die für die Einhaltung der Gütemerkmale Objektivität und Reliabilität mitunter ein entscheidendes Gestaltungsmerkmal sind. Das Ziel bei der Konstruktion von Aufgaben zum Zwecke der Überprüfung der Bildungsstandards im Rahmen eines Sys-

temmonitorings ist die valide Entwicklung von Indikatoren für das Erreichen jeder Kompetenzstufe (Klieme et al., 2003). Dazu gehört u.a. die Klarheit aller Beteiligten über das methodische Vorgehen, der Einsatz einer breiten Aufgabensammlung gekoppelt mit rotierenden Testheftvarianten, der Einsatz entsprechender statistischer Auswertungsmethoden sowie die Einhaltung der lerngruppenübergreifenden Qualitätsmerkmale (Testgütekriterien).

Wenngleich Helmke (2009) auch anführt, dass Aufgaben aus den Bildungsstandards nicht logisch abzuleiten sind, so bilden sie doch das Medium, über das die Schülerinnen und Schüler in eine Performanzsituation gebracht werden, die sie zum Handeln auffordert. Aufgaben stellen gewissermaßen den Operator der Kompetenzmessung dar, der so konstruiert sein sollte, dass er Schülerinnen und Schüler dazu motiviert, sich mit den gestellten Problemsituationen auseinanderzusetzen. Dieser motivationale Aspekt bei der Lösung von Aufgaben ist umso bedeutender, als dass hier eine konkrete Handlung nachgefragt wird, die in ihrer Umsetzung auf Seiten der Lehrkraft diagnostische Schlüsse über den weiteren Kompetenzaufbau zulässt. Die Aufgabenbearbeitung stellt nach Blum (2006) im Mathematikunterricht in jedem Fall die dominierende Schülertätigkeit dar. Deshalb betont er: „...entscheidend sind nicht die Aufgaben selber, sondern ist die Art ihrer Behandlung.“ (Blum, 2006, S. 18). Die Untersuchungen von Zeitler, Heller und Asbrand (2012) lassen den Schluss zu, dass der Einsatz kompetenzorientierter Aufgaben bei einem konstruktivistisch geprägtem Unterrichtsverständnis der Lehrkraft die Adaption des didaktischen Konzepts der Kompetenzorientierung wahrscheinlich macht und dadurch eine Weiterentwicklung des Unterrichts erreicht werden kann.

2.4 Zusammenfassung und zentrale Aussagen

Dieses Kapitel zielte darauf ab, einen kompetenzorientierten Mathematikunterricht zu beschreiben und möglichst klar herauszustellen, welche Änderungen mit der Innovation „Bildungsstandards“ verbunden sind. Dabei lag der Fokus durch die Fachspezifität der Bildungsstandards auf der Darstellung eines kompetenzorientierten Mathematikunterrichts. Die gute theoretische Fundierung des Kompetenzmodells und die Geschwindigkeit der Entwicklungsarbeiten führten dazu, dass „...der Mathematik die Vorreiterrolle bei der Einführung der Bildungsstandards in Deutschland“ zufällt (Prenzel, 2005, S. 262). Die wesentlichen Aussagen des Kapitels stellen sich wie folgt dar:

- Für die Innovation „Bildungsstandards“ lässt sich ein eigenständiger Kompetenzbegriff ausmachen, der von dem Konstrukt der Intelligenz abzugrenzen ist. Als

messbares Merkmal trägt der Kompetenzbegriff die Eigenschaften der Kontextabhängigkeit und der Erlernbarkeit, wodurch von einem kumulativen Kompetenzaufbau auszugehen ist. Kompetenzen werden in Performanzsituationen beobachtbar, aus denen auf eine allgemeine Kompetenz geschlossen wird.

- Die Abbildung von Kompetenzen und deren Aufbau erfolgt durch Kompetenzmodelle, die theoretisch fundiert sind und in Verbindung mit Kompetenzstufenmodellen eine Diagnose sowohl von Aufgabenanforderungen als auch von Schülerfähigkeiten zulassen. Im Fach Mathematik wird eine Einordnung jeder kompetenzorientierten Aufgabe in ein dreidimensionales Modell aus den Dimensionen „Leitidee“, „prozessorientierte Kompetenzen“ und „Anforderungsbereiche“ möglich. Das Kompetenzstufenmodell bildet eine allgemeine mathematische Kompetenz auf fünf Stufen ab, die dem Prinzip des kumulativen Kompetenzaufbaus folgen.
- Auf der Mikroebene des Unterrichts zielt die Innovation „Bildungsstandards“ auf eine Weiterentwicklung der bisher noch vorherrschenden fragend-entwickelnden Unterrichtsgestaltung, mit der eine veränderte Lehrerrolle und die Notwendigkeit eines konstruktivistisch geprägten Unterrichtsverständnis der Lehrkräfte einhergehen. Der Unterricht soll vornehmlich schülerorientiert und effizient gestaltet werden, mit vielfältigen Möglichkeiten zum Kompetenzerwerb und einem kognitiv aktivierenden Anspruch, der durch den Einsatz entsprechender Aufgaben eingelöst wird. Die Leistungen und Einstellungen der Schülerinnen und Schüler verbessern sich im Mathematikunterricht durch eine störungsarme und fehleroffene Lernatmosphäre, in der Fehler zugelassen werden und eine individuelle Anpassung des Lerntempos, der Lerntiefe, des Unterstützungsangebots und Motivationsgrades an die individuellen Lernvoraussetzungen erfolgt. Die Einführung und Anwendung der neuen Aufgabenkultur darf dabei jedoch nicht nur eine Abwechslung zum üblichen Unterrichtsmaterial sein, sondern soll im Verständnis der Innovation zu einer Weiterentwicklung von Unterricht durch Reflexionsprozesse und die Ausrichtung an kumulativen Lernprozessen führen.

Um Lehrkräfte über die Philosophie und die Inhalte der Bildungsstandards zu informieren, liegen bereits vielfältige Materialien vor (z.B. Artelt & Riecke-Baulecke, 2004; Blum, Drüke-Noe, Granzer, Köller, Bremerich-Voß et al., 2010; Hartung, & Köller, 2006; Walther, van den Heuvel-Panhuizen, Granzer & Köller, 2007). Da aber in den ein-

zelen Bundesländern keine einheitlichen Implementationsbemühungen erkennbar sind (KMK, 2005; Krapp, 2004) und kein einheitliches Verständnis der Bildungsstandards vorherrscht (Zeitler et al., 2012), bleibt trotzdem offen, ob und wie Lehrkräfte diese Materialsammlungen nutzen, um im alltäglichen Unterrichtsgeschehen kompetenzorientiert zu arbeiten und die Unterrichtsqualität im Sinne der von Blum (2006) formulierten Komponenten weiterzuentwickeln. Das folgende Kapitel fokussiert daher auf die Möglichkeiten, Erfahrungen und Gelingensbedingungen der Implementation einer Schulreform und wird dafür auch internationale Erkenntnisse in den Blick nehmen.

3 Die Implementation eines standard-basierten Unterrichts

Mit Bezug auf Abschnitt 1.2 wird ein standard-basierter Unterricht als Innovation betrachtet und kann als Gegenstand bestehender Theorien in einem zeitlichen Rahmen dahingehend modelliert werden, dass eine Vorhersage des Übergangs der Innovation zu einem anerkannten, in die Umwelt integrierten Objekt möglich ist. Die Vielzahl bestehender und in der Forschung angewandter Theorien zur Übernahme bzw. Implementation einer Innovation haben gemeinsam, dass sie einen zeitlichen Rahmen abstecken, in dem die Innovation verschiedene Phasen durchläuft, die über eine Entscheidung der Anwender zur Annahme oder Ablehnung der Innovation führt und in einem optimalen Verlauf in der routinierten Anwendung des Objektes gipfelt, womit sich der Begriff Innovation aufhebt. Eine sehr prominente Innovationstheorie legte Rogers bereits 1962 mit der Diffusionstheorie vor, die er stetig weiterentwickelte und von seiner ursprünglichen empirischen Basis aus dem Agrarbereich auf immer mehr Bereiche übertrug. Rogers (2003) untermauert seine Diffusionstheorie mit Beispielen aus dem Technologiebereich, dem sozialen Bereich und dem medizinischen Bereich, womit die Anwendungs- und Erklärungsbreite dieser Theorie deutlich wird. Trotz der starken Bedeutung dieser Theorie und vieler Parallelen im Anwendungsbereich der Bildungsforschung soll sie aus Gründen der Allgemeingültigkeit im folgenden Kapitel nicht fokussiert werden. Auch Hochuli (2011) lehnt Rogers Diffusionsmodell für die Implementation der Bildungsstandards im Schweizer Kontext ab. Die Autorin verweist einerseits auf die Probleme bei der Übertragung der Erkenntnisse aus der Implementation von Agrarprodukten auf den Bildungsbereich und andererseits auf die dem Modell fehlende Steuerung. „Typisch am Diffusionsmodell ist auch der geringe Anteil der Steuerung; ein Verständnis, wonach Neuerungen, wenn sie eine kritische Größe von Nutzenden erreicht haben, „von selbst“ weiter verbreitet werden.

Eine Vorstellung, wie sie für eine gezielte Umsetzung von Bildungsstandards nicht genutzt werden kann.“ (Hochuli, 2011, S. 137).

Spezielle Innovationstheorien im Bildungsbereich wurden u.a. von Euler & Sloane, 1998, Fullan 1982, Little & McLaughlin, 1993, Porter et al., 1988 und Spillane et al. 2002 vorgelegt und durch zahlreiche Studien ergänzt, die unterschiedlichste Aspekte von individuellen Veränderungsfaktoren und auch in Organisationen untersuchen. Berner et al. (2008) beschreiben in einem internationalen Vergleich der Gelingensbedingungen zur Implementation von Bildungsstandards mit Bezug auf Ravitch (2002) zwei Paradigmen, die sich mit je eigenen Thesen im Forschungsbereich der Implementationsforschung gegenüberstehen. Im *professional education paradigm* kommen Theorien des professionellen und organisationalen Lernens zum Tragen, die aus einer konstruktivistischen Perspektive die Prozesse des individuellen und kollektiven ‚sense-making‘ besonders hervorheben (Spillane et al., 2002, Hervorhebungen im Original). Im *policymaker paradigm* verlagerte sich der theoretische Fokus von dem früheren Feld der Innovations- und Change-Theorien zu Theorien der policy-Implementation, die unter Berücksichtigung von Dynamiken und Prozessen den Weg der Innovation durch verschiedene Instanzen des Systems verfolgen. Besondere Bedeutung kommt dabei der Mikroebene des Unterrichts zu. „Dieser neue bzw. zusätzliche Fokus wirkt sich nicht nur auf der Ebene von Konzepten aus, die Implementationsprozesse theoretisch zu fassen versuchen, sondern es hat sich auch gezeigt, dass Evaluationen von in Umsetzung begriffenen Reformen, die unter diesem Blickwinkel durchgeführt werden, ganz andere, d.h. pessimistischere Ergebnisse bezüglich Graden der Implementation ergeben.“ (Berner et al., 2008, S. 219). Im pädagogischen Bereich führte der Fokus auf die Prozesse der Implementation zu einer gesteigerten Sensibilität innerhalb der Theorien im *policymaker paradigm*. Mit der Verabschiedung der Vorstellung, ein Produkt zu implementieren, werden die Effekte jedoch schwerer zu beobachten.

Beide Paradigmen finden ihre Anwendung in den theoretischen Ansätzen zur Implementation einer Innovation, weshalb auch aus beiden Lagern Theorien vorgestellt werden.

Das Kapitel 3.2 beschäftigt sich auf der Mikro- und Mesoebene mit den Faktoren, die eine erfolgreiche Übernahme innovativer Konzepte begünstigen. Hierbei stehen anfangs die relativ stabilen Überzeugungen und Einstellungen der Lehrkräfte im Mittelpunkt, die die Trägheit einer Veränderung im Lehrerhandeln vermuten lassen. Da solche Veränderungen jedoch kein Selbstläufer sind, empfehlen Oelkers und Reusser (2008) die inten-

dierten Unterrichtsveränderungen durch Fortbildungen in die Schulen zu bringen. Angesichts der Notwendigkeit, ca. 660.000 Lehrkräfte auf diesem Weg erreichen zu wollen, zielt der Abschnitt 3.2.2 im Rahmen der Professionalisierung auf die Gestaltung der benötigten Fortbildungen. Es wird untersucht, wie Fortbildungen konzipiert werden sollten, um eine nachhaltige Veränderung der Unterrichtspraxis zu erreichen. Als spezielle Form des Lernens im Beruf wird im folgenden Abschnitt das Konzept der professionellen Lerngemeinschaften vorgestellt, dessen empirische Evidenz erfolgversprechend scheint.

3.1 Innovationstheorien im Anwendungsbereich der Bildungsforschung

Die aktuell große Wissensgrundlage über Innovationsprozesse im Bildungsbereich basiert entscheidend auf den Erkenntnissen aus dem us-amerikanischen Raum. Deutschsprachige Publikationen setzen vielfach dort an und adaptieren die Erkenntnisse auf den untersuchten Kontext (Pant et al., 2008). Dabei stehen eine Vielzahl von Theorien aus der Transfer-, Innovations- und Implementationsforschung zur Verfügung, die die Beantwortung der Frage nach der richtigen Implementationsstrategie erheblich erschweren. Im folgenden Kapitel werden drei Theorien vorgestellt, die ihre spezielle Anwendung im Bereich von Bildungsinnovationen finden. Während Fullan (2007) einen Mehrebenenblick einnimmt, fokussieren Spillane (2004) und Hall & Hord (2006) konkret auf die Aktivitäten der Lehrkräfte.

3.1.1 Fullan's „The new meaning of educational change“

Die Bedeutung von Veränderungen im Bildungsbereich wird von Fullan (2007) in seinem Werk „The new meaning of Educational Change“ beschrieben. Darin konstatiert der Autor einer Innovation eine multidimensionale Bedeutung, die mindestens drei Komponenten oder Dimensionen bei der Implementation eines Programms oder einer Politik umfasst: (1) den Gebrauch von neuen Materialien, (2) die mögliche Anwendung neuer Lehrmethoden bzw. einer neuen Unterrichtsgestaltung und (3) die Änderung in den Überzeugungen oder in den Einstellungen der Lehrkraft. Um eine wirkliche Veränderung im Output zu erzielen, verweist Fullan auf das notwendige Zusammenspiel aller drei Aspekte von Veränderung. Besondere Bedeutung sieht Fullan (2007) darin, dass die Anwender die Innovation verstehen und über sie nachdenken. Die Reflexion über eine Handlung weist auf einen Zusammenhang zwischen der Anwendung und der Übernahme bzw. der Entscheidung des Anwenders, die Innovation anzunehmen hin. „Ultimately it comes down to what is going on in one's head, but the stimulation comes from new experiences that give us something new to think and learn about.“ (Fullan, 2007, S. 41).

Fullan (2007) beschreibt die Veränderung eines Verhaltens (Lehrtätigkeit) bzw. die Anwendung/Übernahme einer Innovation auf einer Zeitachse in drei Phasen, wobei die erste Phase als *Initiierung* oder Konzeption bezeichnet wird. Hier wird die Entscheidung getroffen, ob ein Wechsel bzw. die intendierte Veränderung tatsächlich realisiert oder angegangen werden soll. Die Gründe für die Initiierung einer Innovation sind vielfältig und bestimmen den weiteren Verlauf im Änderungsprozess, wobei auch eine Kombination von Initiierungsgründen möglich ist. So wird eine Innovation, die durch eine Fachinitiative (Lehrercommunity) kombiniert mit einer Orientierung an der Lösung von Problemlagen einen anderen Implementationsprozess zur Folge haben als eine Innovation, die aus bürokratischen Gründen in Kombination mit Öffentlichkeitsdruck entstanden ist. Die Faktoren, die zur Einführung der nationalen Bildungsstandards führten, sind in einer Kombination an der Orientierung zur Lösung einer Problemlage (schlechtes Abschneiden deutscher Schülerinnen und Schüler in internationalen Vergleichsstudien) und einem nachrangigen Öffentlichkeitsdruck (veränderte Anforderungen in der Arbeitswelt erfordern veränderte Fähigkeiten und Kenntnisse, die in der Schulzeit erworben werden) zu sehen. Zur Realisierung einer bundesweiten Wirkung ist als institutioneller Initiator der Bildungsstandards im föderalen Bildungssystem Deutschlands die Kultusministerkonferenz zu betrachten. Der Beschluss zur Einführung der Standards erfolgte ohne eine vorherige ausführliche Diskussion mit den Stakeholdern der Reform, weshalb es sich strategisch betrachtet um eine top-down-initiierte Innovation handelt. Die Analyse bereits erfolgter Implementationsversuche unterschiedlicher Innovationen bezeichnet aus der Verfahrensperspektive zwei klassische Vorgehensweisen: die Übernahme eines fertigen Programms mit autoritativ-interventionistischem Charakter (top-down) oder einen selbstreferenziellen Basisentwicklungs-Ansatz (bottom-up) (Oelkers & Reusser, 2008). Mason et al. (2005) beschreiben top-down-Reformen als theoriegeleitet und forschungsbasiert, wogegen bottom-up-Einführungen die starke Beteiligung von Lehrkräften in der Ausformulierung einer Reform betonen. Gerade die Mitsprache von Lehrkräften ist ein Erfolgsfaktor bei der Übernahme von Reformen. Je geringer dieses Element ausgeprägt ist, desto partieller wird die Reform übernommen (Mason et al., 2005)²². Fullan (2007, S. 82) führt aus, dass der Initiierungsprozess auch von den Personen ignoriert werden kann und der entscheidende Änderungsfaktor im Prozess der Implementation zu sehen ist. „The process of ini-

²² „Top-down reforms tend to apply parts of the research and theory on improving achievement but are generally poorly implemented because they fail to involve teachers in decisions.“ (Mason et al, 2005, S. 356).

tiation can generate meaning or confusion, commitment or alienation, or simply ignorance on the parts of participants and others affected by the change. Poor beginnings can be turned into successes depending on what is done during implementation“. Damit entsteht eine personelle und zeitliche Trennung zwischen Konzeption und Entwicklung der Innovation und deren Umsetzung. Weiterführend ergab sich in den USA daraus die Ursache für frühere gescheiterte Implementationsvorhaben, in dessen Folge anerkannt wurde, dass die Innovation keine Nähe zur Realität des Unterrichts aufwies. Im Nachhinein wurden Praktiker in die Planungen miteinbezogen und der top-down-Ansatz aufgeweicht (Gräsel & Parchmann, 2004). Auch Rogers (2003) diskutiert in seiner Diffusionstheorie top-down oder bottom-up-Verbreitungen einer Innovation. Er beschreibt eine top-down Diffusion in zentralisierten Systemen – von den Experten zu den Nutzern – als zielgerichtet und gesteuert, wogegen in dezentralen Systemen eine Verbreitung durch ein Experimentalverhalten der Anwender, durch Interpretationen und gemeinsamen Erfahrungsaustausch gekennzeichnet ist. Dieser sowohl individuelle als auch kollektive Anpassungsprozess wird von Knapp (1997) als „Inside-out“-Ansatz bezeichnet. Bei diesem Zugang findet häufig ein konstruktivistisches Verständnis von Lernen Anwendung. „Im Zentrum steht nicht die politische Umgebung, sondern der/die individuelle Praktiker/in und die Entwicklung individueller Identität und Kompetenz.“ (Berner & Stolz, 2006, S. 63). Oelkers & Reusser (2008) führen aus, dass komplexe Innovationen kaum je als fertige Programme übernommen werden und als Ergebnis oft nur selektive, oberflächliche und verzerrte Teile der Neuerung implementiert werden. Unter Abwägung der Vor- und Nachteile dieser beiden Strategien erfahren die meisten Implementationsbemühungen aktuell eine Mischform der Regulierung, wie sie auch für die Einführung der Bildungsstandards in Deutschland zu beobachten ist. Damit im Einklang plädiert auch Knapp (2002) im Hinblick auf zukünftige Forschungsdesigns für eine Erweiterung der „Inside-out“-Perspektive, in der sowohl das Umfeld der einzelnen Lehrkraft (professional environment) als auch das Umfeld innerhalb der Schule (community environment) Beachtung finden sollten. Dabei bezieht sich das professionelle Umfeld der Lehrkraft auf die Ausstattung und materielle Unterstützung (auch Fortbildungen) und das Umfeld innerhalb der Schule auf die Personen, mit denen eine gemeinsame Unterrichtsgestaltung erfolgen soll. Als dritten Punkt in zukünftigen Forschungsdesigns führt Knapp (2002) noch die „school and district organisational environment“ an, womit er auf der Exoebene die lokalen Ge-

gegebenheiten einer Schule anspricht (z.B. Lage, finanzielle Ausstattung, Kooperationsmöglichkeiten mit außerschulischen Trägern usw.).

Während die Phase der Initiierung nach Fullan (2007) in Bezug auf die Einführung nationaler Bildungsstandards eine top-down-Entscheidung darstellt, erfährt der Prozess der zweiten Phase, *Implementation*, eine Mischform, die durch einen wechselseitigen Austausch zwischen den Akteuren der Systemebene im Hinblick auf das verordnete Anwenden fertiger Standardpraktiken (*top-down*) gekennzeichnet ist. Fullan (2007) ordnet in diese zweite Phase die ersten Erfahrungen und Versuche ein, die Innovation in der Praxis anzuwenden und bezeichnet sie mit *Implementation* bzw. Erstnutzung. „Implementation consists of the process of putting into practice an idea, program or set of activities and structures new to people attempting or expected to change.“ (Fullan, 20007, S. 84). Der Autor lässt dabei offen, wie die Innovation definiert (initiiert) wurde – als fertiges Produkt mit einheitlicher Umsetzung oder als Entwicklung, die stufenweise durch ihre Anwendung adaptiert wird. In der Phasenlogik des Autors stellt sich am Beginn des Implementationsprozesses die einfache Frage: Was muss sich konkret ändern, wenn die Innovation vollständig implementiert werden soll?²³ Im Kontext von Schule erwartet der Autor Änderungen in den drei anfangs aufgeführten Komponenten eines Programms: Unterrichtsmaterial, Unterrichtsgestaltung und Einstellungen bzw. Verständnis des Curriculums und der Unterrichtsgestaltung. Fullan nennt neun Faktoren, die die Phase der Implementation prägen und die Tiefe der intendierten Veränderung beeinflussen (vgl. Abschnitt 3.3). Damit räumt er gleichzeitig ein, dass die Entscheidung zum Übergang in die nächste Phase – der Verstetigung – auch negativ ausfallen kann. In der Analyse verschiedener Studien zur Einführung von Bildungsinnovation stellt der Autor fest, dass sich die Maßnahmeplanungen für die handlungsintensive Phase der Implementation invers zu den Planungsbemühungen der Initiierungsphase verhalten. Die hohen Erwartungen an die Umsetzung neuer Ideen und das Engagement in der Planungsphase enden oft mit der nicht vollständig durchdachten Implementationsphase: „Commitment to *what should be changed* often varies inversely with knowledge about *how to work through a process of change*.“ (Fullan, 2007, S. 108, Hervorhebungen im Original). Gerade für komplexe schulische Innovationen greifen Innovations- und Change-Theorien zu kurz, da sie nicht ein fertiges Produkt umsetzen, sondern vielmehr an der Produktion eines neuen Produktes

²³ Hierzu Abschnitt 2 mit der Beschreibung dessen, welche Elemente des Mathematikunterrichts einer Änderung bedürfen.

mitwirken. So stellt es sich auch für die Bildungsstandards als Innovation dar. Ihre Implementation erfordert von den betreffenden Personen die Anwendung verschiedener Praktiken, wie es für eine kompetenzorientierte Unterrichtsgestaltung gilt. Folglich wird nicht ein Produkt implementiert, sondern eine Vielzahl von Produkten, die sich gegenseitig bedingen. So führt es auch Fullan in seinen zehn Punkten auf, die er als „do“ und „don’t“ assumptions für eine erfolgreiche Innovationseinführung formulierte: „Assume that changing the culture of institutions is the real agenda, not implementing single innovations. Put another way, when implementing particular innovations, we should always pay attention to whether each institution and the relationships among institutions and individuals are developing or not.“ (Fullan, 2007, S. 124).

Die dritte Phase der Einführung einer Innovation wird von Fullan als *continuation* oder *Verstetigung* bezeichnet. Für diese Phase der Institutionalisierung der Innovation hält der Autor mehrheitlich ernüchternde Ergebnisse bereit. Mit dem Verlust der staatlichen Unterstützung bei der Einführung neuer Programme sinkt auch das Interesse am kontinuierlichen Verstetigen durch Lehrerfortbildungen, Einstellung neuer Kollegen, Mitarbeit an speziellen Projekten usw. Die Gründe für einen Abbruch der Verstetigungsbemühungen sind zumeist gleicher Art derer, die bereits in der Implementationsphase zu Schwierigkeiten führen. Für die Fälle, in denen die Verstetigungsphase erreicht wurde, verweist Fullan auf ernüchternde Ergebnisse zur Nachhaltigkeit der Veränderung. Die Fälle, in denen eine Schulentwicklung erfolgreich gelang, sind in der Unterzahl und zudem können die Entwicklungen, die sich in einer Zeit von fünf bis zehn Jahren vollzogen haben, unter den aktuellen Bedingungen nicht aufrecht erhalten werden: „...furthermore, it appears that the accomplishments are real, but not deep. In other words, even the successful cases cannot be expected to last or to be deep.“ (Fullan, 2007, S. 118).

Fullan bietet mit seinem Werk eine Analyse der Innovationsbemühungen in den USA an, die in einer komplexen Theorie mündet, die so viele Verweise auf individuelle Kontexte bietet, dass sie in ihrer Allgemeingültigkeit bestärkt wird. Die angebotenen drei Phasen zur Einführung einer Innovation sind plausible Entwürfe, wenngleich der Ausblick auf eine erfolgreiche Verstetigung neuer Ideen und Programme eher negativ scheint. Dennoch benennt Fullan wichtige Faktoren, die die Einführung einer Bildungsinnovation wahrscheinlich machen, wenngleich bei der Planung bedacht werden sollte, dass im Implementationsprozess mit Gegenwehr zu rechnen ist, die Idee verändert werden könnte oder die gedachten Maßnahmen die richtigen sind.

3.1.2 Sense-making der Agenten

Spillane (2004) betrachtet die individuellen Voraussetzungen der handelnden Agenten im Implementationsprozess aus einer kognitiven Perspektive. Dabei wird davon ausgegangen, dass neue Erfahrungen und Informationen in einem Assimilationsprozess auf bereits vorhandene Wissensstrukturen aufbauen. Er definiert den Prozess des Sense-Making psychologisch als einen „... active process of interpretation that draws on the sense-maker's experience, knowledge, beliefs, and attitudes.“ (S. 76). An diesem Punkt wird die Theorie Spillanes mit dem Konzept des individual capacity (Abschnitt 3.3.1) verbunden, das für die Übernahme von Innovationen individuelle Kapazitäten in Form von Wissen und Fähigkeiten voraussetzt. Das neue Wissen wird nur sehr träge mit dem bereits existierenden Wissen assimiliert²⁴. Die Konstruktion von Meinungen und die Reaktion der Lehrkräfte auf ausgesandte Stimuli der Politik muss zusätzlich in einer durch Interaktionen gekennzeichneten Umgebung betrachtet werden (Spillane et al., 2002), womit die Untersuchung des Sense-making in situierten Kontexten erfolgt (Berner & Stolz, 2006). Dabei wird immer auf vorgängiges Wissen aufgebaut. Eine besondere Schwierigkeit in der Adaption oder Adoption von neuem Wissen besteht dann, wenn von verschiedenen Seiten oder Institutionen unterschiedliche Botschaften mit teils widersprüchlichen Zielsetzungen ausgesandt werden. Dadurch erfolgt oftmals eine Assimilation neuer Ideen von Unterricht und Lernen, ohne die Kernidee hinter dem Neuen zu verstehen (Oelkers & Reusser, 2008).

Die Theorie des Sense-making ermöglicht eine Erklärung dahingehend, dass dasselbe Reformkonzept von verschiedenen Lehrpersonen unterschiedlich verstanden wird. Da das Verständnis neuer Informationen immer an das bereits vorhandene Wissen geknüpft ist, können die aus der Entwicklungspsychologie bekannten Schemata (Piaget, 1979) nur unter den Bedingungen der Passung im Sinne eines vorhandenen Wissenspools erweitert werden. Schemata sind dabei miteinander verlinkte Wissensstrukturen, die bekannte Konzepte zueinander in Beziehung setzen und somit ein Verständnis des Weltgeschehens erzeugen. Die Haupttätigkeit einer Person im Prozess des Sense-making liegt im Vergleichen von bestehendem und vermeintlich neuem Wissen und im Interpretieren bzw. Dekodieren neuer Informationen. Als Ergebnis bauen Personen ihre eigenen intuitiven Modelle als mentale Repräsentationen darüber auf, wie bestimmte Dinge zusammenhängen und worin deren Handlungsursprünge liegen. Damit verweist das Konzept des

²⁴ „Sense-making tends to be a conserving process.“ (Spillane, 2004, S. 76)

Sense-making auf das auch in der Lehr- Lernforschung bekannte Konzept der subjektiven Theorien (vgl. Abschnitt 3.3.1). Spillane, Reiser & Reimer (2002) verbinden mit diesen intuitiven Modellen auch die Anwendung bzw. Nichtanwendung vorgegebener Unterrichtspraktiken, indem die Lehrkraft den Mehrwert dieser neuen Praktik mit den bestehenden Annahmen darüber, was passieren wird, abgleicht. Die Klarheit entsteht aus der Bestätigung des Bekannten, wenngleich ein neues Verständnis nur durch die Überarbeitung der bestehenden Schemata erzeugt wird (Akkomodation, Spillane, 2004). Ebenso verhält es sich mit den Personen der übergeordneten Ebene. Spillane (2004, S. 79) kann zeigen, dass die politischen Entscheidungsträger (policymakers) keinen Unterschied im neuen mathematischen Problemlösen zu den ihnen bekannten Sachaufgaben erkennen konnten: „For many district policymakers, problem solving was all about linking school mathematics more directly to real-life situations.“. Aus Gründen der Vertrautheit und der Nähe zum bereits Bekannten verbreiteten sie die mit den us-amerikanischen Mathematik- und Naturwissenschaftsstandards verbundenen Lernkonzepte des „hands-on science“ und „integration“ sehr viel häufiger in den Schulen als die Ideen des „constructivist learning“ and „conceptual understanding“. Der Autor sieht die Schwierigkeit der Übernahme neuer Ideen darin begründet, dass Personen erstens neue Ideen als bereits bekannt ansehen und zweitens nur die oberflächlichen Ähnlichkeiten und Merkmale zwischen dem neuen und dem bereits existierendem Wissen wahrnehmen. Die konzeptuellen Elemente der mit der Einführung der Standards bezweckten Änderung des Unterrichtsverständnisses wurden weitestgehend ignoriert. Die Entscheidungsträger assimilierten das neue Wissen ohne die bestehenden Skripte und Schemata zu überarbeiten bzw. neue entstehen zu lassen. Zeitler, Heller und Asbrand (2012) finden diese Form der Wahrnehmung in ihren Untersuchungen zur Implementation der deutschen Bildungsstandards in den habituellen Orientierungen brandenburgischer Lehrkräfte wieder. Die Lehrkräfte begriffen die Bildungsstandards als eine aus DDR-Zeiten bekannte Standardisierung des Unterrichts über den Input und die Vereinheitlichung von Prozessen. „Die Einführung der Bildungsstandards wirkte in der Gruppe „Rhein“ nicht als Anregung für die Weiterentwicklung, sondern als Bestätigung der früheren Praxis standardisierten Unterrichtens.“ (Zeitler et al., 2012, S. 217). Spillanes kognitives Model der Implementation verortet den Grad der Umsetzung des Standard-Gedankens im Unterricht. In 25 Klassen untersuchte der Autor die Unterrichtsgestaltung im Hinblick auf das mit den Standards intendierte Verständnis von Mathematik, wobei er drei Ebenen der Implementation kategorisierte, in die Lehrkräfte eingeordnet

werden können. Dabei bildet die Ebene 3 die größte Nähe zur vorherrschenden Unterrichtsgestaltung im Fach Mathematik ab, indem die Aufgaben eine Rechenprozedur verlangen, die von der Lehrkraft vorgeführt und von den Schülerinnen und Schülern eingeübt und auswendig gelernt wird. In dieser Gruppe verortet Spillane elf Lehrkräfte. Nach einer intensiven Übungsphase wird zum nächsten Problem übergegangen. Wenngleich als Sozialform im Unterricht auch die Gruppenarbeit gewählt werden kann, verbleiben als Problemstellungen doch Aufgaben, die eine eingeübte Rechenprozedur verlangen, ohne über die Zusammenhänge zu anderen Themengebieten anzuregen oder Vernetzungen und Verständnis zu erzeugen. „In this respect, the instruction they experienced overlapped with the mathematics standards. But what counted as mathematics and doing mathematics was firmly grounded in procedural knowledge and computing right answers.” (Spillane, 2004, S. 150).

Ein Unterricht auf der Ebene 1 des Implementationsmodells bietet Aufgaben an, die mathematisches Problemlösen in Alltagssituationen einfordern und bei den Lernerinnen und Lernern ein prinzipielles Verständnis von Mathematik hervorrufen. Die Bearbeitung komplexer Aufgaben, verbunden mit der Aufforderung zur Erklärung von Lösungswegen und der Untersuchung alternativer Vorgehensweisen (Reflexion) wird von einer Lehrkraft initiiert, die durch die Fragetechnik dazu ermuntert, so häufig wie möglich mathematische Konzepte anzuwenden. Der Unterricht von gerade vier Lehrkräften wird dieser Gruppe zugeordnet. Die verbleibenden zehn Lehrkräfte der Ebene 2 in Spillanes Modell der Implementation gestalten ihren Unterricht irgendwo zwischen diesen beiden Entwürfen. Die verwendeten Aufgaben orientieren sich zwar am grundsätzlichen mathematischen Wissen und fordern zum mathematischen Problemlösen auf, die Gespräche im Klassenraum fokussieren sich jedoch auf die Anwendung einer Rechenprozedur. Die Fragetechnik der Lehrkraft ist möglichst eng gestaltet, so dass es für Schülerinnen und Schüler nur eine mögliche richtige Lösung der Aufgabe gibt. „Teachers in the second group infrequently attempted, and rarely managed, to bring to the surface students’ mathematical thinking.” (Spillane, 2004, S. 152).

Wenngleich der Autor die Mehrheit der Lehrkräfte in der Gruppe 1 verortet, geben doch alle an, den Standards besondere Aufmerksamkeit zukommen zu lassen. 21 der 25 Lehrkräfte geben an, mit den Standards „eher“ oder „ganz und gar“ vertraut zu sein. Spillane sucht nach der Ursache für den von vier Lehrkräften der Gruppe 3 auf die Standards abgestimmten Unterricht. Doch weder die Art der Lehrerbildung, noch das verwendete

Curriculum oder das Interesse an Mathematik bieten befriedigende Erklärungen an, da diese Elemente auch in der Gruppe der Lehrkräfte auf Ebene 1 vorzufinden sind. Die drei der vier Lehrkräfte der Gruppe 3 berichten von einer intensiven und zielorientierten Zusammenarbeit im Kollegium. Die vierte Lehrerin stieß in ihrem Kollegium auf Ablehnung, zehrte jedoch von ihrer kürzlich abgeschlossenen Ausbildung, in der die Standards thematisiert wurden. Sie baute ein soziales Netzwerk zum Austausch von Ideen und Erfahrungen außerhalb ihrer Schule auf. Auch die anderen 21 Lehrkräfte engagierten sich, die Standards einzuführen und besuchten Workshops. Im Gegensatz zu den vier Lehrkräften der Gruppe 3 verblieben sie mit ihren Erkenntnissen zumeist allein. „A suburban elementary teacher noted, „There isn’t a daily ‘How do you teach this?’ Or ‘What did you do for this?’ That kind of interaction doesn’t happen.” (Spillane, 2004, S. 160).

Spillanes Theorie des Sense-Making bietet eine Erklärung für unvollständige oder missverständliche Implementationen von Reformvorhaben an, die bei den Wahrnehmungen der Personen ansetzt und auf den durch Piaget bekannt gewordenen kognitiven Anpassungsvorgängen der Assimilation und Akkomodation aufbaut. Der Autor bezieht ebenso die Organisations- und Politikumstände in seine Betrachtungen mit ein. Er macht deutlich, dass das Sense-making der Stakeholder (Lehrkräfte, Schulpolitiker usw.) von diesen Rahmenbedingungen abhängt und nicht ohne diese betrachtet werden kann. Auf individueller Ebene kommt Spillane jedoch zu dem Schluss, dass die Weiterentwicklung des Unterrichts in Relation zum Ausgangsunterricht gesehen werden muss und deshalb bereits die Verwendung anderer Aufgaben einen ersten Schritt der erfolgreichen Implementation ausmacht.

3.1.3 Concerns-based-Adoption-Model

Aus einer langen Forschungstradition zum Änderungsverhalten von Individuen, Organisationen und Systemen extrahieren Hall & Hord (2006) 12 Prinzipien, deren Beachtung bei der Planung und Einführung von Innovationen im Schulsystem einen Zeitgewinn und ein geringeres Problemaufkommen versprechen. Die Prinzipien werden als der Rahmen verstanden, in dem Umbrüche stattfinden und Innovationen eingeführt werden und aus dem die Forschergruppe ihre Annahmen bildet:

1. Veränderung ist ein Prozess, kein Zeitpunkt.
2. Es gibt bedeutende Unterschiede, die die Innovationsentwicklung und – implementation mit sich bringen.
3. Organisationen verändern sich erst, wenn sich die Personen in ihr verändert haben.
4. Innovationen umfassen unterschiedliche Größen und Inhalte.

5. Eingriffe sind als Maßnahmen bzw. Ereignisse zu betrachten, die als Schlüsselkomponenten des Erfolgs dienen.
6. Es erfolgt keine Leistungssteigerung solange keine neuen Praktiken eingeführt wurden.
7. Die Unterstützung der übergeordneten Institution ist essentiell für eine langfristige Änderung.
8. Aufträge (im Sinne einer top-down-Implementation) können funktionieren.
9. Die Schule ist die Haupteinheit der Veränderung.
10. Unterstützung ist eine Teamleistung.
11. Angemessene Eingriffe reduzieren den Veränderungswiderstand und
12. die Kontextbedingungen einer Schule beeinflussen den Änderungsprozess.

Ausgehend von einer grundsätzlichen Schutzhaltung der Individuen gegenüber Veränderungen und den jedermann innewohnenden Assoziationen, die die Veränderung mit sich bringt, betrachten die Wissenschaftler Implementationsprozesse aus verschiedenen Perspektiven und adressieren dabei die System, Organisations- und Teamebene. Die vorangegangene Forschung stellt bei der Betrachtung von Umbruchsituationen für jede Perspektive gut elaborierte Konstrukte und Modelle bereit, die die Autoren referieren und damit die Komplexität einer Untersuchung von Veränderung verdeutlichen. Der Beitrag von Hall und Hord (2006) in der Thematik von Innovationsprozessen mündet im Concerns Based Adoption Model (CBAM), das auf der individuellen Ebene die nötige Unterstützung zur Erreichung der intendierten Ziele identifiziert. Das Model entstand als Konzeption für change facilitators, die einen Innovationsprozess begleiten und unterstützen. Als Concerns definieren die Autoren: „The composite representation of feelings, preoccupation, thought, and consideration given to a particular issue or task (...) To be concerned means to be in a mentally aroused state about something.” (S. 138-139).

Eine umfassend differenzierte Analyse des individuellen Standpunktes innerhalb eines Innovationsprozesses wird durch die Anwendung von drei Instrumenten gewährleistet: die Beschreibung der real vorzufindenden Situation vor Ort durch eine *Innovation Configuration Map* (ICM), die Beschreibung des Verhaltens, das sich aus dem Vorhandensein der Innovation ergibt (*Level of Use*) und die Messung der persönlichen Concerns, die mit der Innovation assoziiert werden. Während die Innovation Configurations und Level of Use als qualitative Interviews durchzuführen sind, wurden die Stages of Concerns in

Form eines Fragebogens entwickelt, der in der internationalen Forschung breite Anwendung findet (van den Berg, 2002; Bitan-Friedlander, Dreyfus & Milgrom, 2004).

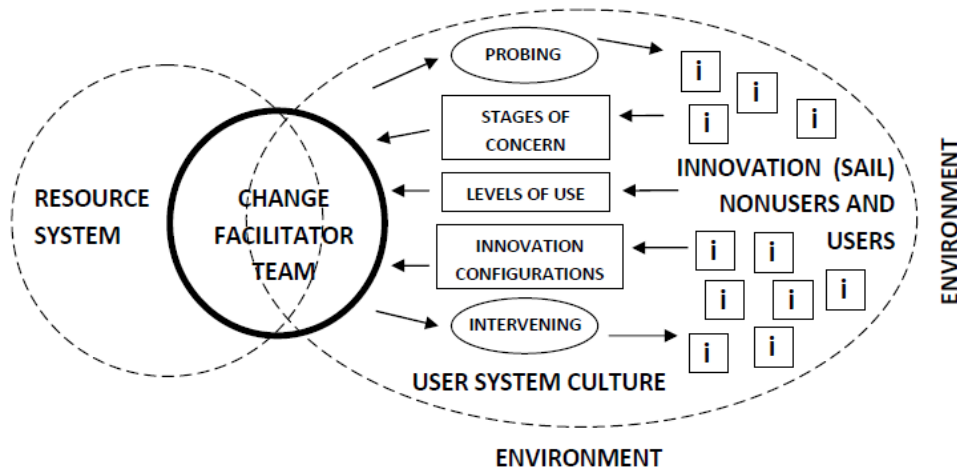


Abbildung 9. Das Concern-Based-Adoption-Model (Hall & Hord, 2006, S. 108).

Der Concern-theoretische Ansatz von Hall und Hord (2006) beschäftigt sich genauer mit der Auseinandersetzung von Lehrkräften mit Innovationen, also mit Gefühlen, Gedanken, Besorgnissen (sog. Concerns) die Personen durchleben, wenn sie sich in einer Umbruchsituation befinden. Dabei wird davon ausgegangen, dass sich die erfolgreiche Übernahme einer Veränderung in mehreren Phasen vollzieht, in denen sich Personen jeweils mit verschiedenen Themenschwerpunkten auseinandersetzen: Wird eine Innovation neu eingeführt, haben Personen demnach zunächst einen *selbstbezogenen Fokus* in ihrer Auseinandersetzung mit dem Neuen. Nachdem die Personen die Möglichkeit hatten, sich mit den neuen Konzepten vertraut zu machen, treten stärker *aufgabenbezogene Concerns* in den Vordergrund. Schließlich setzen sich Personen mit *außenbezogenen Concerns* auseinander.

Im SoC-Fragebogen werden die drei Phasen der Auseinandersetzung nochmals differenziert erfasst, sodass insgesamt sieben Stufen von Concerns unterschieden werden, die als Prozess-Stadien konzeptualisiert werden. Die selbstbezogenen Concerns werden dabei genauer über drei verschiedene Stufen erfasst:

- Auf der Eingangsstufe Kein Bewusstsein (Stufe 0) haben Lehrkräfte keine oder kaum eine Vorstellung von der Innovation und verspüren auch wenig Bedürfnis, sich damit zu beschäftigen.
- Bei der Stufe I Information beginnt eine wenig emotional getönte Auseinandersetzung im Sinne der Informationsbeschaffung.

- Schließlich folgt bei Stufe II Persönliche Betroffenheit ein Fokus auf die persönlich anstehenden Veränderungen und Konsequenzen.

Die aufgabenbezogenen Concerns werden mit einer Phase dargestellt:

In dieser Stufe III Aufgabenmanagement konzentriert sich die Aufmerksamkeit auf die organisatorischen, logistischen und technischen Anforderungen, die mit der Neuerung in- und außerhalb des Klassenzimmers einhergehen.

Die Phase der außenbezogenen Concerns wird wieder in drei Stufen differenziert:

- Zunächst stehen auf der Stufe IV Auswirkung auf Lernende Fragen der Wirksamkeit, der Effizienz oder der negativen Begleiterscheinungen der Innovation für die direkt betroffenen Schüler und Schülerinnen im Mittelpunkt.
- In Stufe V Kooperationsmöglichkeiten verschiebt sich der Fokus der Beschäftigung auf das Bestreben, die neu aufgebauten Handlungsmuster in formalen und informellen Kooperationen mit anderen Lehrkräften zu validieren und zu festigen.
- In der Stufe VI Revision und Optimierung werden schließlich weiterreichende oder alternative Formen der Innovation erprobt.

Idealtypisch verschieben Individuen im Verlauf eines Implementationsprozesses den Fokus ihrer kognitiven Beschäftigung nacheinander durch die sieben Stufen, wobei der Ausgangspunkt innerhalb der ersten drei Stufen liegen sollte. Lehrkräfte würden demnach von den anfänglich selbstbezogenen Concerns, die sich während der ersten Jahre der Innovationsanwendung einstellen, über die aufgabenbezogenen Concerns zum Fokus der außenbezogenen Concerns nach drei bis fünf Jahren bewegen. Der Prozess stoppt jedoch, wenn Lehrkräfte nicht die benötigte Unterstützung erhalten oder eine Innovationsüberfrachtung von politischer Seite stattfindet. Die Lehrkräfte würden dann auf der Stufe der selbstbezogenen Concerns verharren.

Zur Erfassung des Grades der Verwendung von Elementen der Innovation bieten Hall & Hord (2006) ein Verzweigungsinterview an (Level of Use). Im Interview können sich Lehrkräfte auf ebenfalls sieben verschiedenen Stufen der Arbeit mit einer Innovation befinden. Die Stufen 0 bis II werden als Nichtnutzer geclustert und die Stufen III-VI als Nutzer: Lehrkräfte können keine praktischen Schritte zur Umsetzung der Innovation unternehmen (Stufe 0 - *Keine Umsetzung*), sie können beginnen, Informationen zu sammeln (Stufe I - *Orientierung*), sie beginnen die Implementation für sich zu planen (Stufe II - *Vorbereitung*), sie wenden die Innovation rezeptbuchartig an (Stufe III - *Schematische Umsetzung*), Lehrkräfte können mit der Implementation vertraut sein und sie zunächst

„lehrbuchartig“ implementieren und später verfeinern (Stufe IV - *Routine und Verfeinerung*), schließlich können Lehrkräfte bei der Umsetzung mit anderen kooperieren (Stufe V - *Kooperation*) und die Implementation weiterentwickeln und praktisch alternative Umsetzungen explorieren (Stufe VI - *Weiterentwicklung*)²⁵.

Theoretisch postulieren die Wissenschaftler einen Zusammenhang zwischen den Stufen des Fragebogens (SoC) und den erreichten Stufen im Interview. So korrespondieren die aufgabenbezogenen Concerns aus dem SoC mit der Stufe der schematischen Umsetzung im LoU. Empirisch konnte dieser Zusammenhang jedoch nicht nachgewiesen werden.

Die Informationen aus dem SoC geben Hinweise auf kognitive Auseinandersetzung mit der Innovation. Die Daten aus dem Interview (LoU) zeigen den innovationsbegleitenden Personen das Ausmaß der Nutzung. Offen bleibt, welche Teile oder Merkmale der Innovation wie genutzt werden. Zur Schließung dieser Datenlücke entwickelten Hall & Hord (2006) die Innovation Configuration map (IC). An einer solchen Orientierungskarte arbeiten verschiedene Interessengruppen der Innovation. Die Karte beschreibt die intendierten Stationen und Komponenten im Innovationsprozess und führt alternative Wege zur Zielerreichung auf: „The IC map is composed of „word picture“ descriptions of the different operational forms of an innovation or change.“ (Hall & Hord, 2006, S. 116).

Die Anwendung der drei von den Wissenschaftlern entwickelten und vorgestellten Instrumente zur Standortbestimmung einzelner Personen oder Personengruppen in einer Umbruchsituation ermöglicht es den prozessbegleitenden Beteiligten, Interventionen zu planen und durchzuführen, um eine erfolgreiche Implementation zu gewährleisten. Gerade zu Beginn eines Veränderungsprozesses, wenn die selbstbezogenen Concerns überwiegen, wird zu einer vielschichtigen Kommunikation auf vielen verschiedenen Kanälen geraten. Auch das Bewusstmachen des Scheiterns empfehlen die Autoren, wenn sich die Personen, die eine Innovation implementieren sollen, auf der Stufe 1 des SoC (Information) befinden. Andernfalls besteht die Gefahr des Verharrens auf der Stufe 2 der persönlichen Betroffenheit: „When people don’t know what is happening, it is perfectly normal for Stage 2 Personal concerns to become more intense. The less information that is provided, the higher the Stage 2 Personal concerns will be.“ (Hall & Hord, 2006, S. 151).

Die Concern-Theorie nach Hall & Hord (2006) geht zusammenfassend davon aus, dass eine Person im schulischen Umfeld einen ca. 5-jährigen Innovationsprozess durchläuft, der anfangs von selbstbezogenen Gedanken und Gefühlen geprägt ist und mit einer In-

²⁵ Die Übersetzung der Stufen erfolgte in Anlehnung an Pöhlmann et al. (2014).

formationssuche oder Nutzungsvorbereitung einhergeht. Im Verlauf einer begleitenden individuell gestalteten Unterstützung verschieben sich die Befürchtungen (concerns) hin zu Gedanken über die zu erledigenden Aufgaben mit mechanischer Nutzung der Innovation. Als Implementationsende definieren die Wissenschaftler eine kognitive Auseinandersetzung hinsichtlich alternativer Möglichkeiten der Nutzung und einen reflektierten Umgang mit der Innovation.

3.2 Lehrerkognitionen und Potentiale der Veränderung der Unterrichtspraxis in kooperativen Strukturen

Spätestens seit der viel beachteten Metaanalyse von Hattie (2009) ist klar, dass so mancher Lehrer den Unterschied im Schulerfolg ausmacht, d.h. es kommt auf diejenigen Lehrkräfte an, „die bestimmte Unterrichtstätigkeiten im Rahmen passender Curricula anwenden, und die den Lernenden zeigen, wie man in Bezug auf diese Curricula *Denkweisen* und *Strategien* entwickelt.“ (Hervorhebungen im Original, Hattie, 2013, S. 42). Die Publikation betont erneut, dass die Varianz der Schülerleistungen innerhalb der Schulen größer ist als zwischen den Schulen. Mit einer zu 30 Prozent erklärbaren Varianz der Schülerleistungen durch Merkmale, die in der Lehrperson begründet sind, soll im folgenden Kapitel der Blick auf die Faktorenbündel des Lehrerhandels gelenkt werden, die sich für eine Veränderung des Unterrichts im Sinne einer an Standards orientierten Unterrichtsgestaltung herauskristallisiert haben.

Dazu werden im Abschnitt 3.2.1 die theoretischen und empirischen Diskussionen zur Veränderbarkeit von Einstellungen und Überzeugungen von Lehrkräften untersucht, denn essentiell bei der Frage nach der Unterrichtsgestaltung scheinen relativ fest verankerte Schemata über das Lernen zu sein, die sich aus dem Erfahrungswissen aus der Praxis aufbauen.

„Change will always fail until we find some way of developing infrastructures and processes that engage teachers in developing new understandings.“ (Fullan 2001, S. 37)²⁶. Die Veränderung subjektiver Theorien und Überzeugungen scheinen der Schlüssel zum veränderten Lehrerhandeln zu sein, denn letztlich stellt das Lernangebot im Unterricht die naheste Quelle für den Kompetenzerwerb der Schülerinnen und Schüler dar. Im Rahmen der Einführung von Bildungsstandards bleibt dabei zu beachten, dass das Konzept an ca.

²⁶ „Der Wandel wird immer fehlschlagen, bis wir einen Weg finden Infrastrukturen und Prozesse zu entwickeln, die Lehrer dazu veranlassen ein neues Verständnis zu entwickeln.“ (Übersetzung d.Verf.)

660.000 deutsche Lehrkräfte herangetragen werden muss²⁷, die als erfahrene Lehrer über ein breites Repertoire an Wissen, Fähigkeiten Strategien und Erfahrungen verfügen (Lipowsky, 2010). Abschnitt 3.2.2 vertieft daher die Kenntnisse zur Wirksamkeit von Lehrerfortbildungen. Traditionell sollen Schulen und Lehrpersonen auf dieser *intermediären Ebene* des Schulsystems die Möglichkeit haben, intentionsgemäße Handlungen aufzubauen, aufrechtzuerhalten und sozial zu koordinieren (Altrichter, 2010, Hervorhebungen im Original).

DuFour und Eaker definieren einen professionell agierenden Lehrer als Lernender des Lehrens und außerdem als Konsumenten von Forschung (DuFour & Eaker 1998, S. 220)²⁸. Sie betonen die Bedeutung des Einsatzes von Forschungsberichten und neuen Erkenntnissen, um Schulen zu verbessern und die Reflexion über die eigene Unterrichtspraxis anzuregen. Den Raum dafür könnten Lehrerteams innerhalb der Organisation Schule bilden, die auf einem gemeinsamen Interesse und Vertrauen fußen. Wenn Fullan (2007) konstatiert, dass jede erfolgreiche Veränderung sozial basiert erfolgt, ist das Lernen des Einzelnen darüber hinaus auch in der Gemeinschaft des Lehrerkollegiums zu betrachten, wodurch die Bedeutung von professionellen Lerngemeinschaften innerhalb einer Organisation stark zunimmt, da der gegebene Schutzraum für die Lehrkräfte die Motivation birgt, über Unterricht zu sprechen, zu reflektieren und damit weiterzuentwickeln. Seiner Einschätzung nach besitzt die Entwicklung professioneller Lerngemeinschaften ein größeres Entwicklungspotenzial für kontinuierliche Reformen als bisherige Konzeptionen (Fullan 2000, S. 15). Diese Einschätzung wird von vielen us-amerikanischen Schulentwicklern und Schulentwicklungsforschern geteilt (Hord; Senge; Rosenholtz; Mitchell/Sackney; Rolff; DuFour/Eaker). In Kapitel 3.2.3 werden die Charakteristika dieser speziellen Lehrerteams vorgestellt und ihr Potential für die Implementation einer Innovation näher beleuchtet.

3.2.1 Unterrichtliches Handeln auf der Grundlage von Einstellungen, Überzeugungen und subjektive Theorien

Mit dem von Baumert und Kunter (2006) vorgelegten Modell zur Einteilung der professionellen Handlungskompetenz von Lehrkräften (COACTIV) weisen die Autoren den Überzeugungen und Werthaltungen über das Lehren und Lernen neben dem Professions-

²⁷ In Oelkers & Reusser (2008) wird von 800.000 Akteuren gesprochen, als Quelle für die Anzahl von 660.000 Lehrkräften siehe: <http://de.statista.com/statistik/daten/studie/162263/umfrage/anzahl-der-lehrkraefte-nach-schularten/> [Zugriff am 26.04.2014]; ohne Lehrkräfte in Schulkindergärten und Vorklassen.

²⁸ Übersetzung d. Verf. „Professional teachers are students of teaching and consumers of research.“

wissen, den motivationalen Orientierungen und den selbstregulativen Fähigkeiten eine tragende Rolle zu, wenn es darum geht, Lehrerkompetenzen darzustellen. Überzeugungen, Einstellungen und subjektiven Theorien werden - als nicht klar voneinander abgrenzbare Konstrukte - handlungsleitende Funktionen zugeschrieben, die die Unterrichtspraxis einer Lehrkraft bestimmen können und im Rahmen dieser Arbeit die Annahme von Innovationen bzw. Umsetzung von Reformen über die erfahrungsgeleiteten Auffassungen von Lehren und Lernen beeinflussen (Brunner et al., 2006, Dubberke et al., 2008).

In dem von Reusser, Pauli und Elmer (2011) vorgelegten Review zur Systematik der „Berufsbezogenen Überzeugungen von Lehrerinnen und Lehrern“ konstatieren die Autoren die Uneinigkeit und bisher nicht erfolgte Abgrenzung in den Begrifflichkeiten der Einstellungen und Haltungen, motivationalen Orientierungen und Werten²⁹. Dabei lässt sich der in der internationalen Literatur weitestgehend mit *Überzeugungen* übersetzte Begriff der *beliefs* anhand der Merkmale eines intentionalen Gegenstandsbezugs, einer affektiven Aufladung und Wertbindung, eines individuell verinnerlichten Habitus, der Stabilität und Resistenz gegenüber Umstrukturierungen ausweist und einer schwierigen Zugänglichkeit charakterisieren. Die Autoren beschreiben Überzeugungen als *theorieförmige*, quasi-logische Strukturen im Sinne rekonstruierbarer, mehr oder weniger elaborierter semantischer Netzwerke. Durch die Breite der Definition können Überzeugungen „...individueller oder kollektiver Natur, explizit oder eher implizit (intuitiv), fragmentarisch oder sogar widersprüchlich sein oder sich zu personalisierten praktischen (subjektiven) Theorien bzw. zu mehr oder weniger kohärenten, theorieförmigen Handlungs- und Aussagesystemen verbinden.“ (Reusser et al., 2011, S. 478).

Für Kunter und Pohlmann (2009, S. 267) beinhalten die Lehrerüberzeugungen „...Vorstellungen und Annahmen von Lehrkräften über schul- und unterrichtsbezogene Phänomene und Prozesse mit einer bewertenden Komponente.“. Im Fokus dieser Arbeit stehen insbesondere die Erkenntnisse, die die Autoren zum Zusammenhang von Überzeugungen, Einstellungen bzw. subjektiven Theorien und unterrichtlichem Handeln sowie darauf aufbauend auf die Schülerleistungen veröffentlicht haben. Es gilt hierbei, aus den bisher vorliegenden Studien das Potential herauszufiltern, die bestehenden Überzeugungen der Lehrkräfte zum Lehren und Lernen hinsichtlich einer intendierten Entwicklung zu

²⁹Das als „messy construct“ bezeichnete Konzept der Lehrerüberzeugungen unterliegt bis heute keiner klaren Differenzierung. Reusser et al. (2011) führen diesen Umstand teils auf unterschiedliche Forschungsperspektiven zurück, die von unterschiedlichen Grundannahmen ausgehen und in verschiedenen Forschungstraditionen verankert sind.

verändern, um den allgemein akzeptierten Einfluss auf das Lehrerhandeln in die gewünschte Richtung zu lenken. Wenn Kunter und Pohlmann (2009) ausführen, dass die auf verschiedenen Systemebenen angesiedelten Überzeugungen einer Lehrkraft inhaltlich auch durchaus inkongruent sein können, dann wird umso deutlicher, wie differenziert die Überzeugungen im Sinne eines intentionalen Gegenstandsbezugs zu betrachten sind: es ist vorstellbar, dass Lehrkräfte prinzipiell eine positive Einstellung gegenüber reformorientierten Unterrichtsansätzen besitzen, die Umsetzung im eigenen Unterricht mit einer bestimmten Klientel jedoch für nicht machbar erachten. Für diese Arbeit ergibt sich daraus die Möglichkeit, dass Lehrkräfte den Bildungsstandards offen und positiv gegenüberstehen, einen kompetenzorientierten Unterricht aufgrund der Kontextbedingungen im eigenen Handlungsfeld der Schule jedoch nicht realisieren können oder wollen.

Anhand mehrerer Studien wird das Merkmal der *Reflexion* als Voraussetzung einer bewussten Auseinandersetzung mit den eigenen Überzeugungen herauskristallisiert. Ernüchternd zeigt sich jedoch auch, „...dass sich Veränderungen der Überzeugungen im Anschluss an Weiterbildungsmaßnahmen zwar auf die Rhetorik, aber nicht auf das Handeln der Lehrkräfte auswirkten (Forgasz & Leder, 2008).“ (Reusser, Pauli & Elmer, 2011, S. 489). Den größten Erfolg zur Veränderung dieser gegenstandsbezogenen handlungsleitenden Kognitionen (Blömeke et al., 2003) scheinen verschiedene Mentoring- und Trainingsprogramme aus dem Forschungskontext der subjektiven Theorien zu versprechen. Kunter und Pohlmann (2009) sehen im Ansatz des „conceptual change“ eine Erklärungsmöglichkeit für die Stabilität und Veränderungsresistenz von Lehrerüberzeugungen. Demnach lassen sich Überzeugungen am ehesten verändern, wenn die bisher bestehenden Konzepte zur Erklärung von bestimmten Phänomenen nicht ausreichen oder neue Konzepte bereitstehen, die plausibel und erklärungs mächtig sind. Von großer Bedeutung für das professionelle Handeln der Lehrkräfte scheinen die lern- und schulbezogenen Überzeugungen zu sein. Dubberke et al. (2008) untersuchten den Einfluss von lerntheoretischen Überzeugungen auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern³⁰. Grundlegend kann eine Unterscheidung in zwei Lehrertypen vorgenommen werden, die sich in ihrer Auffassung über das Lehren und Lernen unterscheiden. So kann Lernen dabei als das Akkumulieren von Wissensstoff verstanden werden. Die Lehrkraft sieht sich hier als Instruktur und teilt dem Lernenden eine passive, rezeptive Rolle

³⁰ Nach Kunter und Pohlmann (2009, S. 272) beschreiben lerntheoretische Überzeugungen als, „...die Annahmen und Wertvorstellungen, die Lehrende über Lehr-Lernprozesse haben; sie beziehen sich spezifisch auf das jeweilige Fach [...] oder auf Lehren und Lernen i. Allg. [...]“.

zu (Helmke, 2009). Speziell für den Mathematikunterricht fassen die Lehrkräfte Wissen als eine „...statische und unveränderbare Sammlung von algorithmischen Werkzeugen“ auf (Dubberke et al., 2008, S. 194), die von der Lehrkraft zum Lernenden transferiert wird. In einer anderen Auffassung wird Lernen als aktive Konstruktion und Reflexion von Wissen verstanden, in der Problemlösesituationen im Vordergrund stehen und die Lehrkraft bei der mathematischen Aufgabenwahl eher auf ein konzeptuelles Verständnis bei der Lösung der Aufgaben baut (Dubberke et al., 2008; Helmke, 2009; Staub & Stern, 2002). In einer Studie mit 21 us-amerikanischen Mathematiklehrkräften der vierten bis sechsten Jahrgangsstufe zeigen Stipek et al. (2001) im Zusammenhang zwischen „teachers' beliefs and practices“, dass Lehrkräfte mit eher traditionellen Überzeugungen zum Mathematiklernen, mehr Wert auf die Leistungen der Schülerinnen und Schüler legen in Form von richtigen Antworten, Bearbeitungsgeschwindigkeit und guten Noten. Diese Lehrkräfte kontrollieren das Lernen ihrer Schüler stark und lassen ihnen wenig Raum für Individualität und Autonomie, Fehler sollen minimiert werden. Auf der anderen Seite wird Kreativität bei der Lösung mathematischer Probleme eher von Lehrkräften mit einem konstruktivistischen Lehr-Lernverständnis assoziiert. In der Ansicht dieser Lehrkräfte sollten Schülerinnen und Schüler Mathematik eher verstehen als Prozeduren anzuwenden. Schülerinnen und Schüler sollten einen gewissen Grad an Autonomie besitzen und interessieren sich für herausfordernde Aufgaben, auch wenn es nicht um eine Note geht. Lehrkräfte mit einer konstruktivistischen, lernbezogenen Überzeugung denken, dass die Mathematikfähigkeit eines Schülers durchaus veränderbar ist, während die Vertreter der traditionellen Ansicht davon überzeugt sind, dass die Mathematikfähigkeit einer Person relativ stabil und nur schwer veränderbar ist. Diese Überzeugung wird hinsichtlich des Lehrerhandelns assoziiert mit einer Leistungs Betonung (schnelles Rechnen und richtige Antworten). Die Empfehlungen der Autoren zum Verändern der Überzeugungen von Mathematiklehrkräften sind konsistent mit den Empfehlungen anderer Wissenschaftler: Lehrkräfte sollten neue Praktiken in ihrem Unterricht ausprobieren und diese Stunden in einem kooperativen Setting reflektieren: „... - to move back and forth among a variety of setting to learn about new instructional strategies, to try them out in their own classrooms, and to reflect on what they observed in a collaborative setting.“ (Stipek et al., 2001, S. 225). Bensen und Berkemeyer (2011) führen hinsichtlich des Lehrerhandelns jedoch aus, dass die in den Köpfen der Lehrkräfte verankerten persönlichen Unterrichtsskripts und Überzeugungen aufgrund ihrer relativen Stabilität nur schwer veränderbar sind und eine

tatsächliche Umsetzung von Innovationen nur dann erfolgt, wenn diese eine Verbindung zu ebendiesen Skripts und Überzeugungen zulassen. Veränderungen erfordern daher eine Rekontextualisierung der Überzeugungen der Lehrkräfte hinsichtlich neuer Rahmenbedingungen. Da beide Konstrukte als handlungsleitende Kognitionen beschrieben werden (Blömeke, 2003) bleibt jedoch fraglich, ob eine Veränderung des Unterrichtsskriptes durch Ausprobieren o.ä. die Veränderung der Überzeugungen bewirkt oder ob der Wirkungszusammenhang eher umgekehrt eintritt.

In der Thematik der Bildungsstandards liefert die rekonstruktive Studie von Zeitler et al. (2012) erste Hinweise auf die Anschlussfähigkeit der Vorstellungen vom Lernen der Schülerinnen und Schüler an die Umsetzung eines kompetenzorientierten Unterrichts. Heller (2012) kann darin zeigen, dass Gymnasiallehrkräfte eher einem instruktivistischen (traditionellen) Unterrichtsverständnis folgen und Lehrkräfte an integrierten Gesamtschulen ein konstruktivistisches Lehr-Lern-Verständnis aufweisen. Gründe sieht die Autorin in den unterschiedlichen professionellen Bezugssystemen der beiden Gruppen von Lehrkräften. Während das Lehren und Lernen am Gymnasium eher durch Fachlichkeit charakterisiert werden kann, wird in integrierten Gesamtschulen eher pädagogisch unterrichtet: „Die instruktivistische Vorstellung von Unterricht [...] steht dagegen dem Zulassen oder gar Fördern und Unterstützen der eigenständigen Konstruktionsprozesse der Schülerinnen und Schüler entgegen.“ (Heller, 2012, S. 200). Gymnasiallehrkräfte stehen vor der kaum zu bewältigenden Herausforderung, die Anforderungen eines kompetenzorientierten Unterrichts, wie etwa die Förderung unterschiedlicher Lernprozesse und –wege, umzusetzen. Aufgrund der unklaren Abgrenzung der Begrifflichkeiten im Bereich der Lehrerkognitionen werden Überzeugungen auch teilweise als Einstellungen, Vorstellungen oder Auffassungen bezeichnet (Reusser et al., 2011). In Anlehnung an die bisherigen Ausführungen werden für die Einführung der Bildungsstandards zwei Überlegungen relevant: die Lehrkräfte (1) brauchen eine größtmögliche Klarheit über die Konzeption und Zielsetzung der Bildungsstandards und sie (2) sollten vom Nutzen, den die Standards bringen überzeugt sein. „Darüber hinaus ist von zentraler Bedeutung, dass die Bildungsstandards von den Lehrkräften als nützliches Instrumentarium angesehen werden,... (Freudenthaler & Specht, 2006, S. 8). In einer Evaluationsstudie von Beer (2007) zu den Einstellungen von 110 österreichischen Lehrkräfte gegenüber den Bildungsstandards werden anhand von 23 Items sowohl positive als auch negative Aspekte der Einführung von Bildungsstandards erfragt. In der Auswertung zeigt sich, dass die Mittelwerte der Items zu den positiven

Aspekten im Wesentlichen nahe um den Skalenmittelwert von 3 schwanken. Der Autor zieht daraus eine kritische Bilanz. Eine erfolgreiche Implementation würde mit Werten assoziiert, die näher am positiven Pol der Zustimmungsskala liegen. Insgesamt stehen die Lehrkräfte den positiven Argumenten zur Einführung von Bildungsstandards eher verhalten gegenüber. Die Mittelwerte der relevanten Items lagen unter dem theoretischen Mittelwert der fünfstufigen Antwortskala. Demgegenüber wurden die Items zu den problematischen Aspekten der Bildungsstandards von den Lehrkräften differenzierter bewertet und streuten viel stärker. Freudenthaler und Specht (2006) folgerten daraus eine hohe Unsicherheit bei den Lehrkräften und empfahlen, diese mit Informationsveranstaltungen und Fortbildungen zu verringern. Im Implementationsbericht des in Österreich federführenden Bundesinstituts für Bildungsforschung, Innovation und Entwicklung des österreichischen Schulwesens (Bifie) wird die nach der evaluierten Pilotphase ausgesprochene Empfehlung bestätigt. Aus den Daten der 2009 befragten österreichischen Lehrkräfte lässt sich schließen, dass die positive Einstellung gegenüber den Bildungsstandards steigt, umso eher sich die Befragten informiert fühlen. Hinsichtlich des unterrichtlichen Handelns wird von Grillitsch (2010) konstatiert, dass die Lehrkräfte umso weniger die Bildungsstandards bei der Unterrichtsplanung und –gestaltung heranziehen, je weniger sie sich informiert fühlen. Insgesamt erfolgt nur eine vereinzelte Verwendung der Bildungsstandards im Unterricht von den Lehrkräften. Rund ein Drittel gibt an, die Bildungsstandards gar nicht für die Unterrichtsplanung und –gestaltung heranzuziehen. Informations- und Unterstützungsmaßnahmen werden für die Entwicklung einer positiven Einstellung gegenüber den Bildungsstandards auch weiterhin als förderlich bzw. als wesentlicher Bestandteil des Implementationsprozesses betrachtet. Die benötigte Akzeptanz der Bildungsreform wird von der Autorin daran festgemacht, inwiefern Lehrkräfte die Bildungsstandards als praxistauglich erleben (Grillitsch, 2010).

In Deutschland deuten die Ergebnisse von Pant et al. (2008a, b) auf eine ähnliche Einstellung von Lehrkräften gegenüber den nationalen Bildungsstandards hin. Von den Autoren wurde mithilfe des standardisierten *Stages of Concern*-Befragungsinstrumentes (Abschnitt 3.1.3) eine große Gruppe von so genannten „Kooperierern“ identifiziert, die sich durch eine individuelle Auseinandersetzung mit den Bildungsstandards auszeichnen, sich aber gleichzeitig mehr Informationen über die Implikationen der Innovation wünschen. Der im typischen „M-Profil“ abgebildete Lehrertypus möchte im Sinne einer Kooperation im Team gern am Wissen der Kolleginnen und Kollegen teilhaben, ohne eine handlungs-

orientierte Zusammenarbeit zu präferieren. Dennoch wurde ein Viertel der befragten Lehrkräfte³¹ als „Einzelkämpfer“ identifiziert, die hinsichtlich einer Innovation sehr selbstbezogen agieren und kaum im Kollegium kooperieren. Noch 21 Prozent der Sekundarstufenlehrkräfte wurden als „informationssuchende Kooperierer ohne Aufgabenfokus“ eingestuft. Diese Lehrkräfte suchen verstärkt nach Informationen, ohne jedoch die Auswirkungen für das weitere Handeln in den Blick zu nehmen. Pant et al. (2008a) kommen zu dem Schluss, dass die Mehrheit der Lehrkräfte die durch Bildungsstandards implizierten Unterrichtsformen entweder noch nicht erkennen oder sie als theoretische Anforderung verstehen. Diese erste systematische Untersuchung zur Implementation der Bildungsstandards weist damit darauf hin, dass Lehrkräfte den Bildungsstandards zwar aufgeschlossen gegenüberstehen, dass sie sich aber noch ungenügend über die Innovation informiert fühlen und kompetenzorientierter Unterricht noch nicht zum Schulalltag der Lehrkräfte gehört. Idealerweise sollten sich Lehrkräfte dann in einigen Jahren mit der konkreten Umsetzung der Standards kognitiv beschäftigen und sich stärker mit den bisher noch wenig ausgeprägten praktischen Anforderungen der Implementation auseinandersetzen, d.h. „...die Aufmerksamkeit auf die organisatorischen, logistischen und operativen Anforderungen, die mit der Neuerung in- und außerhalb des Klassenzimmers einhergehen, wie z.B. erforderliche Änderungen im Zeitmanagement durch neue Lehrformen“ zu konzentrieren (Pant et al., 2008, S. 831).

Speziell für Mathematiklehrkräfte in der Sekundarstufe I legen Pöhlmann et al. (2014) Ergebnisse einer Interventionsstudie vor, die hinsichtlich der Auseinandersetzung mit der Reform „Bildungsstandards“ die Notwendigkeit einer begleitenden Professionalisierungsmaßnahme im Implementationsprozess bestätigt. Die Lehrkräfte in der Interventionsgruppe (mit prozessbezogener Begleitung) setzten sich im Laufe eines Jahres viel stärker mit dem Konzept der Bildungsstandards auseinander als Lehrkräfte, die keine Unterstützung bei der Implementation erhielten. Weiterhin gaben die Lehrkräfte an, vermehrt kompetenzorientierte Tätigkeiten im Unterricht einzufordern, womit sich auch das Unterrichtshandeln der Lehrkräfte in der Interventionsgruppe veränderte. Die Autoren führen diese Veränderungen auf die Prozessbegleitung zurück, die durch die Beratung und Unterstützung so genannter Setkoordinatorinnen gegeben war³².

³¹ Die gezogene Stichprobe war hinsichtlich ihrer Schulzusammensetzung bundesweit repräsentativ (Pant et al., 2008a)

³² Der Untersuchung liegt die gleiche Stichprobe zu Grunde wie sie auch Teil dieser Arbeit ist, auch Abschnitt 5.1.1.

Demgegenüber sieht Fullan (2007) die Möglichkeit der Veränderbarkeit von Einstellungen und Überzeugungen als gegenstandsbezogene Kognitionen, deren Wichtigkeit in einem Prozess der Einführung einer Innovation als essentiell betrachtet wird, als sehr schwierig an. Er konstatiert einschränkend: „(1) behaviors and emotions change before beliefs – we need to act in a new way before we get insights and feelings related to new beliefs – and (2) the size and prettiness of the planning document are inversely related to the amount and quality of action, and in turn to student achievement (Reeves, 2006), and (3) shared vision or ownership (which is unquestionably necessary for success) is more of an *outcome* of a quality change process than it is a *precondition* for success.” (Fullan, 2007, S. 41, Hervorhebungen im Original). Die damit entstandene Anzweiflung der Wirkrichtung im Sinne der Reihenfolge der Änderungen führt zu einem anhaltenden Forschungsdesiderat. Es ist kaum zu klären, da beiderseits plausibel, ob zuerst die Einstellungen und Überzeugungen der Lehrkräfte verändert werden müssen, um ihr Lehrerhandeln zu verändern oder sollten Lehrkräfte die neuen Techniken zuerst ausprobieren, also Handeln, bevor sich ihre Überzeugungen und Einstellungen ändern. Die Vertreter der Lehrerkognitionsforschung äußerten sich dahingehend, dass dem manifesten Lehrerhandeln kognitiv gesteuerte Entscheidungsprozesse unmittelbar vorausgehen (Reinisch, 2009, S. 39).

3.2.2 Professionalisierung im Kontext von Unterrichtsentwicklung

Lipowsky (2010) versteht unter Lehrerkognitionen im erweiterten Sinne einerseits Überzeugungen und subjektive Theorien, andererseits das fachliche, fachdidaktische und pädagogisch-psychologische und diagnostische Wissen von Lehrkräften. „Insofern lässt sich erfolgreiches Lernen im Rahmen von Professionalisierungsmaßnahmen zum einen an der Veränderung beruflicher Überzeugungen, subjektiver Theorien und deklarativer Wissensanteile der Lehrer festmachen, zum anderen an der Erweiterung und Flexibilisierung unterrichtlichen Handelns.“ (Lipowsky, 2011, S. 399). Die Modellierung professioneller Handlungskompetenz, wie sie Baumert und Kunter (2006) differenziert anhand eines heuristischen Rahmenmodells darstellen, ist mit Blick auf die Veränderung des Lehrerhandelns daher nicht ausreichend zielführend (wenngleich die Autoren ein viel beachtetes Modell zur Abbildung der professionellen Handlungskompetenz von Lehrkräften vorlegen). Angesichts der großen Masse an erfahrenen Lehrkräften bieten die Konzeptionen zum Kompetenzaufbau in der Lehrerbildung zwar einen Ansatzpunkt, geeigneter erscheinen jedoch Theorien der Trainingsforschung bzw. Weiterbildung. Hattie (2009)

stellt im Anschluss an die Präsentation der wirkungsrelevanten Faktoren von Unterricht im Hinblick auf verbesserte Schülerleistungen die Frage, warum sich Lehrpersonen nicht ändern können bzw. warum Lehrpersonen den statistischen Belegen nicht vertrauen. Aus den Metaanalysen zu erfolgreichem Lernen sieht er eine tiefgehende Änderung des Unterrichtens dann angestoßen, wenn Schulen sich zu einer evidenzbasierten Auffassung bewegen könnten. Er stellt fest, dass Lehrpersonen die Ursachen für schlechte Schülerleistungen einerseits bei den Lernenden suchen und andererseits an beispielhaften Anekdoten festhalten, die ihren Unterricht bestätigen. In dem Bedürfnis der Lehrkräfte nach Sicherheit, Kontrolle und Einfachheit suchen sie eher nach Beispielen, die ihre aktuellen Methoden bestätigen und sehr selten nach empirischen Belegen, warum etwas nicht funktionierte. Ein gemeinsames Verständnis von Unterricht basiert daher auf Erzählungen: „So often teachers depended on „war stories“, personal experiences, and a reliance on their own experience to justify their personal preferences. If this swapping of war stories is the closest teachers come to professional conversations, the picture is bleak for the messages in this book about teachers needing to share evidence about their teaching with their colleagues.” (Hattie, 2009, S.252). Die Lösung zur Aufgabe der Resistenz gegenüber Unterrichtsinnovationen sieht der Autor unter Bezug auf Little (2007) in der Möglichkeit, allgemeine Prinzipien für die Praxis zu erstellen, deren Verankerung sich sowohl in einem konzeptuellen Rahmen als auch in den Besonderheiten der persönlichen Erfahrungen wiederfindet. Die Erstellung solcher Prinzipien erfolgt durch die gegenseitige Befragung, die Wiederholung von Situationen, das Sammeln empirischer Belege in den Erzählungen und dem Anbieten und Revidieren von Interpretationen. Als entscheidenden Faktor zur Erreichung einer solchen Praxis bzw. zur Überwindung der Widerstände gegenüber veränderten Konzepten von Lehren und Lernen betont der Autor die Bedeutung der professionellen Weiterbildung. Lipowsky (2011) sieht berufliches Lernen im Rahmen von Weiterbildungsangeboten theoretisch durch personenbezogene Voraussetzungen, durch Faktoren des Schulkontextes und durch Merkmale und Komponenten der Fortbildung selbst erklärbar. Die Interaktionen modelliert er in einer Art Angebots-Nutzungs-Modell, das ausgehend von einem Weiterbildungsangebot eine unterschiedliche Wahrnehmung, Nutzung und Weiterverarbeitung der Lerngelegenheiten bei den Teilnehmenden annimmt und in einer differentiellen Wirkung der Fortbildung auf die teilnehmenden Lehrpersonen mündet. Gräsel et al. (2006) führen aus, dass es bei der Weiterqualifizierung von Lehrkräften im Sinne einer weitreichenden Veränderung des Unterrichts (wie es mit der Ein-

führung der Bildungsstandards für viele Lehrkräfte der Fall ist) erforderlich ist, die bestehenden Handlungsrouinen und möglicherweise auch tief sitzenden Überzeugungen über Lehr-Lernprozesse zu verändern. Aus der Forschung können die Merkmale der Langfristigkeit, Situierung und Anregung zur Kooperation als wirksame Konzipierungsfaktoren von Fortbildungen bestimmt werden. Dabei machen Gräsel et al. (2006) deutlich, dass kurzfristige Fortbildungen ein Inselwissen gegenüber existierenden Handlungsrouinen und –erfahrungen produzieren. Im Gegensatz dazu ermöglichen langfristig angelegte Fortbildungen, die neuen Inhalte zu erproben und systematisch auf die eigene Unterrichtssituation zu beziehen. Die Situierung erfolgt durch die Lehrkräfte. Fortbildungsinhalte müssen von den Teilnehmern auf den individuellen Unterricht adaptiert werden. Das gelingt, je näher die Inhalte am Lehr-Lerngeschehen liegen und je höher sich der Bezug zum Unterricht darstellt. Das situierte Lernen umfasst dabei über die eigene Erprobung hinaus das Reflektieren anderer Unterrichtssituationen, wobei die Entwicklung von Optionen und Handlungsalternativen für diese Situation mit der zusätzlichen Reflexion darüber entscheidend sind (Schreder, 2007). Kooperatives Arbeiten rahmt den Prozess der Situierung, indem Lehrkräfte zur Verbesserung ihres Unterrichts in einer diskursiven Auseinandersetzung mit anderen bestehende Handlungsrouinen verändern. Es wird angenommen, dass die kommunikative Aushandlung von neu umzusetzenden Inhalten und Maßnahmen zu einer verstärkten Akzeptanz an einer Schule führt und damit die Realisierung der Veränderung fördert. In einer Studie zur Anregung der Kooperation stellte die Forschergruppe in einer ersten Phase fest, dass Lehrkräfte nur schwer einen Zugang zu kooperierendem Arbeitsverhalten finden. In Schulen mit einer bereits bestehenden Kooperationsstruktur erfolgte eher ein fortbildungsbezogener Austausch bzw. eine Zusammenarbeit als in Schulen ohne eine Kooperationskultur. Auch die Leistungen der Schülerinnen und Schüler erwiesen sich als deutlich höher.

Lipowsky (2010) macht für die Untersuchung der Wirksamkeit von Lehrerfortbildungen den Vorschlag einer Verortung nach der Reichweite der Wirkung, wodurch eine Kategorisierung der Fortbildungsziele auf vier Ebenen erfolgt. Auf der Ebene 1 werden die Reaktionen und Einschätzungen der Lehrkräfte dargestellt. Lipowsky (2004) fasst in dieser Ebene die Aussagen zusammen, die den Transfer der Inhalte in die Unterrichtspraxis fördern. So wird mehrheitlich auf den Unterrichtsbezug und die Möglichkeit zum Informationsaustausch mit anschließender Nachbetreuung verwiesen, die die Akzeptanz und den selbstberichteten Kompetenzgewinn begünstigen. Der Autor zieht dabei auch Ergebnisse

aus einer Teilstudie des Projektes „Chemie im Kontext“ (CHiK) heran, in der die Abhängigkeit der Teilnahmemotivation der Lehrpersonen von der erlebten Autonomie und der wahrgenommenen Relevanz nachgewiesen werden konnte. Die Nützlichkeit und Relevanz eines Trainings stellen in verschiedenen Studien der Trainingsforschung auch die Elemente dar, die einen Wissenszuwachs bzw. ein verändertes Verhalten im Beruf bestätigen. Aufgrund der doch sehr schwachen Zusammenhänge zwischen den Reaktionen und Einschätzungen der Lehrkräfte und einem professionellen Kompetenzaufbau sieht Lipowsky (2010) eher eine indirekte Wirkung der Akzeptanz einer Fortbildung auf den Transferprozess über das Engagement. Die Ebene 2 wird mit der Erweiterung der Lehrerkognitionen überschrieben. Die auf die Veränderung der Lehrerkognitionen und Überzeugungen gerichtete Wirksamkeit ist empirisch in einigen Studien nachzuweisen. So konnten die unterrichtsbezogenen Überzeugungen von Lehrkräften in den Studien von Gärtner (2007) und Krammer et al. (2010) im Hinblick auf die gewünschte konstruktivistische Unterrichtshaltung verändert werden. In der Studie von Krammer et al. (2010) berichteten die Lehrkräfte nach Abschluss der auf die Gestaltung eines kognitiv aktivierenden Unterrichts angelegten Fortbildung, dass sie über mehr Wissen hinsichtlich der Gestaltung eines schülerorientierten, kognitiv aktivierenden Unterrichts verfügen. Die Nutzung und Analyse eigener Unterrichtsvideos wurde von den Teilnehmern sehr geschätzt, wenngleich der Einfluss externer Faktoren auf das Unterrichtshandeln der teilnehmenden Lehrkräfte nicht isoliert werden konnte und die Kausalität der Wirksamkeit der Fortbildung dadurch geschwächt wurde. Dennoch betrachtet auch Lipowsky (2010) die Konfrontation der Lehrkräfte mit Unterrichtsvideos als ein Mittel, vertieftes Nachdenken über die eigene Praxis anzuregen, da Dissonanzen zwischen den eigenen Überzeugungen und der Wirkung auf der Schülerseite sichtbar werden. An dieser Stelle findet auch die umgekehrte Wirkrichtung ihren Stellenwert: erst durch das Ausprobieren neuer Unterrichtspraktiken und die erkennbare Wirkung auf Seiten der Schülerinnen und Schüler werden Lehrkräfte dazu angeregt, ihre Einstellungen und Überzeugungen nachhaltig und dauerhaft zu verändern.

Die Ebene 3 „Unterrichtspraktisches Handeln“ verfolgt die Kausalkette, dass Veränderungen auf der Schülerebene durch Veränderungen im Unterricht stattfinden: „...unterrichtliches Lehrerhandeln ist die proximalste aller Lehrervariablen und wenn es zu Veränderungen auf Seiten der Schüler/innen gekommen ist, dürften diese demzufolge am ehesten mit Veränderungen im Unterricht und im unterrichtlichen Handeln der Lehr-

personen zu erklären sein.“ (Lipowsky, 2010, S. 58). Als bedeutende Studie führt der Autor eine Längsschnittuntersuchung von Desimone et al. (2002) an, in der die untersuchten Lehrkräfte angeben, eine tatsächliche Änderung der Unterrichtspraxis erfahren zu haben, und zwar unabhängig davon, wie sie vorher unterrichtet haben. Ein starkes Ausmaß an kognitiv aktivierendem Unterricht wird demzufolge durch die Gelegenheiten zur gegenseitigen Hospitation und das Fokussieren auf das Schülerlernen erreicht. Weitere Studien berichten aus Sicht der Schüler über einen klareren und stärker strukturierten Unterricht, der nach einer Intervention erreicht wurde (Wackermann, 2008). Ebene 4 nimmt die am weitesten entfernte Wirkung von Lehrerfortbildungen in den Blick. Auf dieser Ebene der „Effekte auf Schüler/innen“ referiert Lipowsky Studien, in denen ein Zusammenhang zwischen Veränderungen auf Seiten der Schülerinnen und Schüler und der Lehrerfortbildung abgebildet wird. So wird für die im Experimentaldesign angelegten Studien gezeigt, dass ein differenziertes Feedback zu Hausaufgaben sowie die bei Mathematiklehrkräften verankerten sozio-konstruktivistischen Konzepte zu erhöhten Schülerleistungen führten. Weitere empirische Evidenz für die Wirkung von Fortbildungen auf das Schülerlernen bzw. die Schülerleistung stellen die im Abschnitt 3.3.2 referierten Projekte SINUS und CHiK dar.

Aus dem Literaturreview schlussfolgert Lipowsky, dass wirksame Fortbildungen eher fachdidaktisch eng angelegt sein sollten und die Lernprozesse der Schülerinnen und Schüler in den Vordergrund stellen, wobei mit einem Fokus auf ihre fachbezogenen Konzepte, Vorstellungen und Denkweisen das diagnostische Wissen der Lehrkräfte vertieft wird. Als weiteren Erfolgsfaktor wird die Dissonanz zwischen den eigenen Erwartungen und Überzeugungen einerseits und der eigenen unterrichtlichen Praxis bzw. deren Wirkung auf die Schülerinnen und Schüler andererseits betrachtet. Erlebte Widersprüche führen dabei zu einem eigenen „conceptual change“, der berufliches Lernen auslöst. Die von anderen Wissenschaftlern aufgeführte Langfristigkeit einer Fortbildung kommt bei Lipowsky durch die Empfehlung der Kombination von Reflexions- und handlungspraktischen Erprobungsphasen zum Ausdruck.

Lipowsky (2010, S. 65) formuliert auf der Basis der Forschungsbefunde aus der Trainingsforschung eine vorsichtig positive Annahme darüber, „...dass Fortbildungen unter günstigen Bedingungen sehr wohl Wirkungen auf Schüler/innen entfalten können, wenn gleich zu beachten ist, dass es sich in vielen Studien um selektierte, freiwillige Lehrer-

stichproben handelt, bei denen von einer besonderen Motivation und einem besonderen Engagement für Schule und Unterricht ausgegangen werden muss.“.

Einschränkend sieht Altrichter (2010) für systemverändernde Reformen die Gefahr des Widerstands der Lehrkräfte. Sie stimmen zwar auf der allgemeinen Ebene einer Reform zu, geht es jedoch um ihren konkreten Unterricht, werden die Lehrkräfte skeptisch. Eine wirksame, intentionsgemäße Veränderung der Handlungspraxis sieht Altrichter (2010, S. 27) in einem Typ von Fortbildung, „...in dem Wissensaufnahme, Arbeit an den normativen Implikationen in einer unterstützenden Gruppe, Erprobung praktischer Implikationen und Weiterentwicklung der Arbeitsbedingungen in einem längerfristigen Konzept verbunden werden.“. Ernüchternd führen Bonsen und Berkemeyer (2011) schließlich aus den Befragungen zur Rückmeldung der Studien MARKUS (Helmke, 2004), QUASUM (Peek, 2004) und VERA (Groß, Ophoff, Hosenfeld & Koch, 2006), dass Lehrkräfte eine Veränderung ihres Unterrichts nicht in Betracht ziehen. Die Rückmeldungen werden zwar zur Kenntnis genommen und auch als verständlich und ausreichend empfunden, einen Anstoß zur Unterrichtsentwicklung geben sie jedoch nur minimal. Altrichter weist zudem darauf hin, dass sich das traditionelle Professionskonzept geändert hat. Die notwendige Wissensbasis wird quantitativ erweitert, da sich neue Anforderungen durch die aktuellen Schulentwicklungsprozesse stellen. „Während ‚Kooperation‘ und ‚Reflexivität‘ in traditionellen Professionskonzepten wenig Bedeutung auf der Wissensebene haben, werden sie in neueren Professionskonzepten als zentrale Voraussetzungen zum Aufbau einer fallspezifischen Wissensbasis angesehen“ (Altrichter 2000, S. 156). Im Rahmen der österreichischen „Bildungsstandards-Politik“ (Hervorhebung im Original, Altrichter, 2008) resümiert der Autor hinsichtlich der Implementationsproblematik, dass Fortbildungen weniger eine allgemeine Höherqualifizierung fokussieren, sondern vielmehr die Rolle eines „intermediären Unterstützungssystems“ für die intendierten Veränderungen der Systemsteuerung übernehmen sollten. Damit werden die Institutionen der Fortbildung stärker und zielorientierter als bisher in die staatlichen Transformationsinteressen eingebunden.

Speziell im Zusammenhang mit der Implementation von Innovationen im Schulsystem betonen Fussangel, Rürup und Gräsel (2010) die Bedeutung von Lerngemeinschaften, denen dann eine tragende Rolle zugeschrieben wird, wenn langfristige und nachhaltige Veränderungen des Unterrichts erzielt werden sollen. Der folgende Abschnitt geht daher vertiefend auf die Struktur und Besonderheiten von Lerngemeinschaften ein und arbeitet ihren Mehrwert im Rahmen von Implementationsprozessen heraus.

3.2.3 Professionelle Lerngemeinschaften

Wenngleich der Begriff der Professionellen Lerngemeinschaften darauf hinweist, dass es sich um eine Betrachtung von Charakteristika auf Gruppenebene handelt, macht Hord (1997) in Anlehnung an Fullan (1993, 1994) darauf aufmerksam, dass es das Individuum ist, das ein System oder eine Organisation ändert und damit die einzelne Person die vielversprechendste Komponente im systemischen Wandel ist. Die Fähigkeit zur Zusammenarbeit mit anderen ist damit auch eine Voraussetzung des Individuums, der in der gemeinsamen Wirkung der Gruppe kooperierender Lehrkräfte eine hohe Bedeutung für das erfolgreiche Schülerlernen zugeschrieben wird. Rosenholtz (1989) stellte den Bezug zwischen Lehrkräften in kooperativen Netzwerken oder Gemeinschaften und erhöhtem Schülerlernen erstmals her. Sie fand heraus, dass Lehrkräfte, die sich in ihrem eigenen Lern- und Lehrprozess unterstützt fühlen, sich stärker gegenüber ihrer Tätigkeit verpflichten und sich mit ihr identifizieren und damit effektiver unterrichten als andere. Diese verstärkte Selbstwahrnehmung eines effektiven Lehrerhandelns führt zu einer höheren Wahrscheinlichkeit, neue Unterrichtskonzepte auszuprobieren und anzunehmen und überträgt sie zusätzlich auf andere³³. Das Professionswissen von Lehrkräften war für Rosenholtz (1989) ein wichtiger Bestandteil im Konzept der professionellen Lerngemeinschaften (PLG), wenngleich der Begriff „Professional Learning Community“ erst durch spätere Studien geprägt wurde (Bryk et al., 1999). Stoll et al. (2006, S. 225) betonen im Hinblick auf das Konzept der PLG: „At the heart of the concept, however, is the notion of community. The focus is not just on individual teachers’ professional learning but of professional learning within a community context – a community of learners, and the notion of collective learning.“. Die von Rosenholtz 1989 vorgelegte Studie „Teachers’ Workplace“ gilt als Ausgangspunkt und empirische Grundlage für die darauf folgende Forschung im Bereich der Schul- und Unterrichtsentwicklung. In den noch relativ jungen Forschungsbeiträgen schließen Bonsen und Rolff (2006) auf ein gemeinsames Verständnis der in diesem Bereich veröffentlichten Autoren, PLGen als wirksamen Kontext für Schulverbesserung zu betrachten. Sie arbeiten mit einer Begriffsdefinition, in der die Gemeinschaft als eine Gruppe von Menschen zu verstehen ist, „...die durch gemeinsames Fühlen, Streben und Urteilen verbunden ist. Sie sind personenzentriert und befriedigen Bedürfnisse wie Vertrauen, Fürsorge, Anteilnahme, Besorgtheit sowie Bindung, Verpflichtung und Verbind-

³³ „Further, Rosenholtz found that teachers with a strong sense of their own efficacy were more likely to adopt new classroom behaviors and that a strong sense of efficacy encouraged teachers to stay in the profession.“ (Hord, 1997, S. 12).

lichkeit.“ (Bonsen & Rolff, 2006, S. 169). Professionell bedeutet bei den Wissenschaftlern eine qualifizierte Ausbildung und Orientierung an hohen Standards der Berufsausübung und ein Interesse an Weiterbildung.

Alle (zumeist us-amerikanischen) Autoren verbinden die Vorstellung von „Lehrern als Lerner“ mit dem Begriff der PLG. Als Schlüsselwerte werden in den PLGen die Hilfe-Kultur und Fehlertoleranz betrachtet, wobei die Fehler eine Chance zum Lernen darstellen. Auch Hall & Hord (2006) schreiben in ihrer change-Theorie (Abschnitt 3.1.3) den PLGen als Form der Schulkultur eine tragende Rolle bei der Implementation von Schulreformen zu. Mit Bezug auf einen früheren Beitrag von Hord (1997) werden fünf Dimensionen von Professionellen Lerngemeinschaften identifiziert: (1) shared values and visions, (2) collective learning and application, (3) supportive and shared leadership, (4) supportive conditions und (5) shared personal practice. Die Dimensionen weisen auf eine Abgrenzung von den in den Schulen etablierten Lehrerteams hin, indem sie über die formale Bindung durch Fächer oder Jahrgangsstufen hinausgehen und durch eine emotionale Verbundenheit gekennzeichnet sind. PLGen sollen „das, was wir schon lange machen“ um ganz neue Tätigkeiten erweitern bzw. das Ausüben alter Tätigkeiten um neue Methoden ergänzen. Die gemeinsame Arbeit soll Lehrkräfte dazu befähigen, überholte Tendenzen im Lehrerhandeln zu überwinden und „neue Dinge im alten Licht zu sehen“, d.h. trotz unveränderter äußerer Umstände (Schulsystem, Raumsituation, Ausstattung usw.) die Chancen und Möglichkeiten zu einem Wandel und zu Innovationen zu suchen und zu erkennen (Mitchell & Sackney, 2000, S. 15). Die kollegiale Zusammenarbeit ermöglicht der Lehrkraft einen Wechsel aus der Rolle als Einzelkämpfer in die eines aktiven und engagierten Mitglieds des Kollegiums. Hord (1997) schafft mit ihren Dimensionen eine Vorstellung von Schule, in der

- eine verbindliche, gemeinsame Vorstellung vom Lernen in der Schule existiert und auf die Bezug genommen werden kann,
- die Lehrkräfte einer oder mehrerer Schulen gemeinsam und kontinuierlich lernen und die Ergebnisse auf die Schülerinnen und Schüler übertragen,
- sich der Schulleiter ebenfalls als Lerner sieht und die Meinungen der Lehrkräfte in einen demokratischen Entscheidungsprozess einfließen lässt,

- unterstützende strukturelle und individuelle Bedingungen³⁴ zur Diskussion und Umsetzung von Entscheidungen, neuer Unterrichtskonzepte oder problemorientierter Lösungen geschaffen werden und
- der Unterricht von Kollegen vertrauensvoll (nicht evaluativ) hospitiert und reflektiert wird.

In der deutschsprachigen Literatur berichten Bonsen & Rolff (2006) von einer Pilotstudie zum Ausbaubestand von PLGen, in der sie sich an den Kriterien von Newman (1994) und Kruse et al. (1995) orientieren. Im Gegensatz zu Hord formulieren die Autoren teilweise sehr unterschiedliche Merkmale, die eine PLG ausmachen:

1. Reflektierender Dialog
2. De-Privatisierung der Praxis
3. Gemeinsamer Fokus auf Schüler-Lernen
4. Zusammenarbeit
5. Gemeinsam geteilte Normen und Werte.

Die mit der De-Privatisierung angestrebte Öffnung des Unterrichts (zunächst für Kollegen) vollzieht sich nach Hord (1997) zumeist als letzte Entwicklungsdimension. Die Möglichkeit, über unterrichtsbedingte Unsicherheiten zu reden und neue Wege der Praxis durch den Dialog mit Kolleginnen und Kollegen zu erfahren, setzt ein hohes Maß an Vertrauen voraus. Die Berufspraxis der Lehrkräfte ist jedoch durch Autonomie und Isoliertheit gekennzeichnet, was nach DuFour und Eaker (1998) eine der größten Barrikaden bei der Bildung von professionellen Lerngemeinschaften darstellt.

Fullan (2007, S. 52) konstatiert in seiner change-Theorie, dass jede erfolgreiche Veränderung auf sozialer Zusammenarbeit fußt, die vorher noch nicht gegeben war, womit er die notwendige Bildung von PLGen ebenfalls betont: „All successful change initiatives develop collaboration where there was none before.“. Als Konsequenz auf die durch Studien bewiesene größere Varianz des Lehrerhandelns innerhalb einer Schule als zwischen den Schulen, fokussiert er die Bildung von PLGen, in denen die Lehrkräfte sich gegenseitig hospitieren und mit der Schulleitung an der fortlaufenden Unterrichtsentwicklung zusammenarbeiten, um ein Mehr eines gemeinsamen Verständnisses und einer gemeinsamen Verpflichtung gegenüber dem inhaltlichen Gegenstand der PLG zu erlangen.

³⁴ Hall & Hord (2006) unterscheiden hier in “physical conditions”, womit die Zeit für Treffen gemeint ist genauso wie die Bereitstellung von Ressourcen, und “people capacities“, die durch eine positive Einstellung gegenüber Veränderung, Lernen und Feedback beschrieben werden.

Bonsen & Rolff (2006) adaptierten das von Newman et al. (1996) eingesetzte Befragungsinstrument, das die fünf Merkmale einer PLG operationalisiert, für eine Untersuchung in 124 Schulen zur Wahrnehmung von PLGen durch Lehrkräfte. Der unbefriedigende Rücklauf der Fragebögen und die für eine Interpretation auf Schulebene sinnvolle Entscheidung, eine Schule erst mit fünf ausgefüllten Fragebögen in die Analyse aufzunehmen, führte zu einer Datengrundlage, die sich auf $n=35$ Schulen und $n=199$ Lehrkräfte minimierte. Im Ergebnis bestätigen die Autoren das mehrdimensionale Konzept der PLG und die Notwendigkeit, es differenzierter zu erfassen. Sie kommen zu dem Schluss, dass die Wirksamkeit von PLGen über den Begriff der Kooperation hinausgeht: „Es wäre falsch, das Konzept der PLG einfach auf den Begriff der Lehrerverkooperation zu verkürzen. Emotionale und reflexionsbezogene Komponenten müssen hinzukommen.“ (Bonsen & Rolff, 2006, S. 179). Aus den anfänglich explorativen- und nach erfolgter Korrektur konfirmatorischen Faktoranalysen der Items zu den fünf Merkmalen einer PLG, schlugen sie die Umformulierung des Merkmals „gemeinsam geteilte Normen und Werte“ in „gemeinsame handlungsleitende Ziele“ vor. Für die Schulentwicklungsarbeit im nationalen Kontext deutscher Schulen empfehlen die Autoren zunächst eine Klärung der bereits bestehenden innerschulischen Arbeitsstrukturen sowie die Formulierung von Zielgruppen für die Arbeit in PLGen. Nutzbare Strukturen werden dabei in Fachgruppen, Klassenteams und kompletten Jahrgangsgroups gesehen. PLGen würden genau wie die formal bestehenden Lehrerteams eine Ebene innerhalb der Organisation „Schule“ bilden, die zwischen der Einzelschule als Ganzes und den autonom arbeitenden Lehrkräften liegt. Diese „intermediären Organisationen“ (Bonsen, 2005) stellen einen klaren Unterrichtsbezug in den fach- oder lerngruppenspezifischen PLGen her, der die gemeinsamen Lernaktivitäten der Gruppe erleichtert. Die Isolation der Lehrkraft wird deutlich reduziert, es erfolgt eine Identifikation mit den Zielen und Visionen der Schule und es kommt zu einer geteilten Verantwortung für die Gesamtentwicklung der Schülerinnen und Schüler. Das Potential von PLGen beschreiben die Autoren als eine Synthese von Unterrichts- und Personalentwicklung. Hord (1997) konnte für die PLGen folgende Effekte empirisch belegen:

- erfolgreiches Lernen von Unterrichtstechniken und Erweiterung des professionellen Wissens,
- ein tieferes Verständnis des Unterrichtsstoffes und der eigenen Rolle als Lehrkraft in der Unterstützung der Schülerinnen und Schüler bei der Erreichung der erwarteten Leistungsziele,

- eine höhere Wahrscheinlichkeit, dass die Lehrkräfte über aktuelle fachliche Innovationen informiert sind,
- eine höhere Berufszufriedenheit und weniger Fehltage der Lehrkräfte,
- eine bedeutend höhere Kapazität, sich auf die Bedürfnisse der Schülerinnen und Schüler einzustellen und den eigenen Unterricht entsprechend zu adaptieren,
- eine höhere Motivation an bedeutsamen und nachhaltigen Veränderungen mitzuwirken sowie eine höhere Wahrscheinlichkeit fundamentale und systematische Veränderungen mitzutragen. (Bonsen, 2005).

Die Effekte bestätigen die von Bryk, Camburn & Louis (1999) postulierte Aussage, dass „Professional Communities“ die Schlüssel-Kapazität und notwendige Bedingung auf organisationaler Ebene sind, um Personal- und Unterrichtsentwicklung zu ermöglichen. Mit anderen Worten bilden die auf organisationaler Ebene geschaffenen PLGen den Raum, eine Haltungsänderung der Lehrkräfte hin zu einem professionellen Streben nach ständiger Verbesserung zu erlangen, wobei der Fokus auf das Schülerlernen gelegt wird. Fullan beschreibt die Rolle der professionellen Lerngemeinschaften im Kontext der Unterrichtsentwicklung mit der Entwicklung von „Beurteilungsfähigkeit“ („assessment literate“, Fullan, 2007, S. 142). Allerdings meint Fullan hier eine komplexe Art der Beurteilung:

1. Die Fähigkeit, Schülerleistungsdaten und Schülerergebnisse zu untersuchen und sie kritisch zu betrachten.
2. Die Fähigkeit, gemäß dieser kritischen Betrachtung zu agieren, indem Entwicklungspläne für Unterricht und Schule entworfen werden, um die für eine verbesserte Leistung nötigen Veränderungen zu bewirken.
3. Die Fähigkeit von Lehrern, effektive Mitspieler auf dem Feld der Verantwortlichkeit zu sein, indem sie Daten zur Leistungsfähigkeit der Schule proaktiv und offen gegenüberstehen, und in der Lage sind sich in der Debatte um Nutzen und Missbrauch von Leistungserhebungsdaten in einer Ära der großen Tests behaupten zu können. (Fullan, 2001, S.142).

Damit stellt Fullan eine Verbindung zwischen den PLGen, dem Lehrerlernen und den Schülerleistungen her, indem er in Anlehnung an Rosenholtz (1989) darlegt, dass in den learning-enriched schools eine Untersuchung erfolgt, wie gut die Schülerinnen und Schüler in den Leistungstests abschneiden, die Ergebnisse dann auf den Unterricht bezogen werden und daraus Entwicklungen entstehen. Um möglichst effektive Planungen zu erstellen, vereint eine PLG das berufliche Know-how ihrer Mitglieder und treibt so die

Verbesserung zunächst des Unterrichts voran. Neue Methoden der Evaluation und Leistungsbewertung erhalten einen höheren Stellenwert. Durch die Arbeit mit diesen Methoden „im Kleinen“ erhalten Lehrkräfte eine Wissensbasis, die es ihnen ermöglicht, Jahrgangsstufentests oder internationalen Leistungstests kritisch zu begegnen und deren Ergebnisse zu interpretieren.

Für eine nachhaltige Etablierung von PLGen mahnt Rolff (2001) die große Bedeutung der Schulleitungen an, die die Räume für Unterrichtsentwicklung schaffen müssen:

„Wichtig ist, dass die PLGs durch Schulleitungen unterstützt werden, symbolisch wie organisatorisch. (...) Vor allem müsste die Schulleitung die stundenplantechnischen Voraussetzungen treffen und auch Entlastungen schaffen, indem sie z.B. zulässt, dass hin und wieder eine 6. oder 7. Stunde „abgehängt“ wird, zu der dann die PLG tagen kann.“ (Rolff 2001, S. 4).

3.3 Förderliche Faktoren im Prozess der Implementation einer Bildungsinnovation

Die Implementation einer Bildungsinnovation ist ein Vorhaben, das aufgrund einer Mehrebenenstruktur des Bildungssystems die unterschiedlichsten Akteure verpflichtet, die Intentionen der Neuerung zu akzeptieren und umzusetzen. Die Bildungsinnovation der Bildungsstandards zielt auf den Output von Unterricht in Form von Lernergebnissen, während die Innovation der Kompetenzorientierung auf den Unterricht selbst abzielt. Die Mikroebene des Unterrichts mit seinen Akteuren sind somit die Hauptzielgruppen der beiden Innovationen. Da die Leistung und das Verhalten beider Zielgruppen – sowohl der Lehrkräfte als auch der Schülerinnen und Schüler – von anderen Akteuren beeinflusst werden, liegt ein passendes „Rezept“ für die Implementation bislang nicht vor. Im Folgenden sollen daher Erfahrungen und Forschungsergebnisse aus anderen Projekten referiert werden, die eine Implementation wahrscheinlich machen.

3.3.1 Die Genese von Faktoren aus theoretischer Sicht

Mit Blick auf die aktuelle Situation in Deutschland stützt sich der Implementationsprozess weitgehend auf Theorien des professionellen und organisationalen Lernens. Diese Theorien mit dem Schwerpunkt „Lernen“ beziehen sich auf Personen und Organisationen. Bei der Schaffung der Voraussetzungen für anhaltendes und nachhaltiges Lernen werden die zur Veränderung notwendigen Ressourcen und Unterstützungsangebote in den Blick genommen (Oelkers & Reusser, 2008). Im Sinne eines persönlichen als auch organisationalen Kapazitätsaufbaus sind die Konzepte des professionellen und organisationa-

len Lernens an die des capacity-building anschlussfähig. Aus einer Lernperspektive heraus wird capacity-building von vielen Forschern eher als eine Fähigkeit betrachtet, die es den Akteuren ermöglicht, Reformen angemessen zu begegnen und umzusetzen. Little (1999) sieht in der teacher capacity eine Schlüsselrolle im Umsetzungsgrad einer systemischen Reform. Das Konzept wirkt seit der zweiten Bildungsreformwelle in den USA mehr als eine Leitlinie bzw. als Implementationsprinzip, da der ersten Welle als Antwort auf den 1983 von Ravitch verfassten Bericht „A Nation at Risk“ (National Commission on Excellence in Education, 1983) hauptsächlich top-down-eingeführte Maßnahmen wie eine Erhöhung der Standards und Regulierungen, ein erhöhtes Lehrergehalt und verlängerte Schultage folgten, jedoch keinesfalls capacity aufgebaut wurde oder der Unterricht veränderte wurde (Desimone, 2003). Newman, King und Youngs (2000, S. 261) definieren auf Schulebene capacity als „the collective power of the full staff to improve student achievement schoolwide“.

In der Reflexion eines 10 Jahre andauernden Prozesses der Sicherung und Verbesserung von Schulqualität sieht Stoll (2009) das Konzept des capacity buildings mehr als ein holistisches Konzept; in einem Entwicklungsprozess ändert sich die Bedeutung von capacity von einer anfänglichen drei-Kategorien-Betrachtung (personell, interpersonell und organisationell, Mitchell & Sackney, 2000) zu einer sieben-Kategorien-Betrachtung (Massel, 1998). Auf der Unterrichtsebene werden die Kategorien Wissen, Fähigkeiten und Fertigkeiten von Lehrkräften, Schülermotivation und Lehr-, Lernmaterialien für Schüler und Lehrer ausgemacht. Auf der Schulebene werden diese ergänzt durch die Faktoren „Anzahl der unterstützenden Personen im Klassenzimmer“, „Anzahl und Qualität der Sozialbeziehungen“, „Ausstattung und Organisation und Vernetzung der Schulressourcen“. Berner et al. (2008) spricht von „der Notwendigkeit von Inputs unter Bedingungen der Output-Orientierung“. Bei der Durchführung systemischer Standard-Reformen wurden durch Lehrerfortbildung, Schulausstattung, Lehrmaterialien usw. Rahmenbedingungen geschaffen. Der Aufbau solcher Kapazitäten verstärkt sich, je stärker sich die Innovation oder die Politik von den bestehenden Praktiken unterscheidet (Cohen & Barnes, 1993a). Bezogen auf den persönlichen Lernprozess und professionellen Kapazitätsaufbau geht Fullan (2007, S. 59) sogar noch weiter indem er konstatiert: „Capacity building first, and judgment second – because this is what will motivate more people. Learning in context and learning every day are the keys.“. Erst die Erfahrungen des eigenen Erlebens führen zu einer Änderung der Ansichten und Überzeugungen. Folglich sollten Lehrkräfte erst

eine kompetenzorientierte Unterrichtsgestaltung ausprobieren und erste Erfahrungen sammeln, bevor sich tatsächlich ihre Einstellung gegenüber den Bildungsstandards ändert. Auch Oelkers und Reusser (2008, S. 261) verweisen aus einer Akteur-Perspektive auf die durch Forschungsbefunde gestützte Wirksamkeit des professionellen Lernens. „Veränderungen in der Schul- und Unterrichtspraxis setzen schulbezogen-organisationalen, letztlich jedoch immer individuelle Veränderungen – und damit Lernen – voraus.“ Sie machen fünf Gruppen von Bedingungen für eine wirksame Innovationssteuerung aus, die sie auch für die Implementation als zentral halten:

- Bedingungen, die die Innovation selbst betreffen
- Bedingungen, die sich auf die individuellen Lehrpersonen beziehen
- Bedingungen, die im Zusammenhang mit dem Arbeitsfeld der einzelnen Schule stehen
- Bedingungen, die den regulativen Rahmen und die unterstützenden Kontexte der Reform betreffen
- Bedingungen, die sich auf die konkrete Umsetzung der Innovation beziehen.

Für alle fünf Gruppen von Bedingungen machen Oelkers und Reusser (2008, S. 260/261) konkrete Ausführungen, die Prinzipien und Maßnahmen identifizieren, „...die für eine nachhaltige, den Unterricht und das Lernen von Schülerinnen und Schülern erreichende pädagogische Qualitätsentwicklung als zentral gelten können.“. Mit der Betonung von *stärker autonomen Schulen* und der Qualitätsverbesserung der *Lehrerfortbildung* empfehlen die Autoren zwei Schwerpunktmaßnahmen, die schrittweise angestrebt werden sollten (Hervorhebungen im Original). Hier wird deutlich, dass neben den Schulen als Ganzes besonders die einzelnen Lehrkräfte als Träger der Implementation betrachtet werden. Das konstruktive Sense-Making, das Spillane, Reiser und Reimer (2002) beschreiben als die Konstruktion von Sinn und damit Interpretationsunterschiede der Lehrkräfte bei den politischen Botschaften meinen, wird jedoch von drei Faktoren maßgeblich beeinflusst:

- vorgängiges Wissen und eingespielte Praktiken,
- der soziale Kontext,
- das Design und die Repräsentation der Politik.

Wenn folglich Lehrerfortbildungen als Ansatzpunkt gelingender Implementationsprozesse gesehen werden, dann beschreibt Sense-Making den Prozess der Interpretation der Politiksignale, beruhend auf der Wissensbasis, den Überzeugungen und den Ansichten einer

Person³⁵, der in einen sozialen Kontext eingebettet ist und durch die Art der Kommunikation der Reform ausgelöst werden. Der soziale Kontext wird hier auf den Mikrobereich des Arbeitskontextes beschränkt, womit die Zugehörigkeit zu einer Profession, aber auch die Kommunikation in Kollegien und Netzwerken gemeint ist. Entscheidend für das Sense-Making der Lehrkräfte scheint die Präsentation der politischen Ziele zu sein. In der Art von Textdokumenten wie Standards, Testinstrumenten oder Informations- und Weiterbildungsunterlagen, die in kurzer Form die Ziele und den Zweck der Reform beschreiben, ohne die zugrunde liegenden Ideen und Konzepte näher auszuführen, scheint problematisch. Nach den Ausführungen von Oelkers und Reusser (2008) hat sich gezeigt, dass Lehrkräfte eher eine Innovation implementieren, je spezifischer sie transportiert wird. Damit wird auf die Qualität des Materials, der Informationen, der professionellen Entwicklung, der Leitung und der Instruktion hingewiesen, die es Lehrkräften ermöglichen soll, das neue Konzept bzw. die Kernidee im Sinne eines positiven Sense-Making zu verstehen. Allerdings zeigt sich ein Unterschied zwischen schwächeren und exzellenten Lehrkräften. Ersteren fällt die Implementation leichter, wenn anschauliches Material, konkrete und beispielbezogene Instruktionen beiseite stehen, während Letztere stärker auf Autonomie ansprechen. Für diese Gruppe von Lehrkräften gilt es also, die Spezifität eines neuen Konzeptes gegenüber den Befürchtungen einer Deprofessionalisierung transparent zu machen. Neben der Spezifität lassen sich weitere Faktoren ausmachen, die die positive Entwicklung einer Innovation wahrscheinlich machen:

- Konsistenz, im Sinne einer möglichst hohen Übereinstimmung von Schulreformstrategien mit den Strategien der Politik;
- Autorität, im Sinne einer selbstverständlichen Interaktion zwischen den Trägern der Reform, dem Grad der Mitbestimmung bei Entscheidungen und der Schaffung von Netzwerken;
- Durchsetzungsmacht, im Sinne staatlicher Anreize, wie z.B. monetäre Unterstützung;
- Stabilität in Bezug auf die Schulleitung, den Lehrkörper, die Administration und das politische Umfeld (Oelkers und Reusser, 2008).

Diese von Porter et al. (1988) bestimmten Kriterien beschreiben eine Policy Attributes Theory, die Desimone (2002) auf die Analyse erfolgreicher Implementationsmodelle in

³⁵ "Sense making is not a simple decoding of the policy message; in general, the process of comprehension is an active process of interpretation that draws on the individual's rich knowledge base of understandings, beliefs, and attitudes." (Spillane, Reiser und Reimer, 2002, S. 391).

den USA anwendete. Danach ist die Institutionalisierung einer Politik wahrscheinlicher, je überzeugender ihre Autorität im Sinne einer rechtlichen Verankerung bzw. ihre Übereinstimmung mit sozialen Normen ist, als die Durchsetzungsmacht mittels monetärer Anreize, da diese nur befristet anhalten. Desimone (2002, S. 470) hält fest: „Findings show that while each attribute contributes to implementation, specificity is related to implementation fidelity, power is related to immediate effects, and authority, consistency, and stability seem to be the driving forces of long-lasting change.”.

Gräsel (2010) strukturiert die Faktoren zur erfolgreichen Implementation einer Bildungsinnovation in vier Merkmalsgruppen, in denen jeweils verschiedene theoretische und empirische Aspekte zum Tragen kommen. So bezieht sie in der Gruppe der (1) Merkmale der Innovation selbst die Erkenntnisse aus Rogers Diffusionstheorie (2003) mit ein und nennt damit fünf Merkmale, die „leicht transferierbare Innovationen“ aufweisen:

1. von den Lehrkräften wird ein relativer Vorteil bei der Innovation gegenüber der bestehenden Praxis erkannt,
 2. die Innovation ist kompatibel mit bestehenden Überzeugungen und Werten der Anwender,
 3. eine wenig komplexe Innovation wird schneller übernommen,
 4. eine schrittweise Einführung wirkt weniger bedrohlich und
 5. schnelle Sichtbarkeit erster positiver Ergebnisse führt zu einer beschleunigten Übernahme (Gräsel, 2010; Rogers, 2003).
- Die (2) Merkmale der Lehrerinnen und Lehrer werden in der zweiten Gruppe benannt, worin sich die Autorin hauptsächlich auf den concerns-based-adoption-model-Ansatz stützt (Abschnitt 3.1.3). Mit den (3) Merkmalen der Einzelschule betont sie besonders die Rolle der Schulleitung und die Kooperation im Kollegium. Auf die Umweltfaktoren und die Transferunterstützung wird explizit in der vierten Merkmalsgruppe eingegangen. So werden hier Faktoren wie die Stabilität personeller Konstellationen und die Innovationsdichte benannt. Auch das Vorhandensein von schulübergreifenden Netzwerken und das Angebot an Lehrerfortbildungen finden in dieser Gruppe ihren Raum. Gräsel (2010) gibt damit einen strukturierten Überblick über die Faktoren, die die Verbreitung oder Weitergabe von Transferiertem beeinflussen.

3.3.2 Nachhaltige Implementation erfolgreicher Projekte

In Anlehnung an die von Oelkers und Reusser (2008) und Gräsel (2010) postulierten Merkmale einer erfolgreichen Implementation schulischer Innovationen wird nachfolgend angeknüpft. Die Erfahrungen aus zwei Programmen im deutschsprachigen Raum sollen

beispielhaft dafür Pate stehen, wie die zumeist in den USA entwickelten Implementations-theorien auf den deutschsprachigen Raum übertragen werden können. So wird der Erfolg von professionellen Lerngemeinschaften durch die Anwendung im Projekt „Chemie im Kontext“ (CHiK) bestätigt. Als zentrales Element einer so genannten symbiotischen Implementationsstrategie wurde CHiK von Anfang an als Umsetzung einer Unterrichtskonzeption mit Hilfe einer Implementationsstrategie geplant, die die stetige Zusammenarbeit von Wissenschaftlern und Praktikern in schulübergreifenden Lerngemeinschaften forciert. Die Verbreitung des SINUS-Programms andererseits steht Pate für ein Konzept, das die im Abschnitt 3.2.2 dargestellten Erkenntnisse zur Konzeption von Fortbildungen aufgreift und weiterentwickelt. Die Zusammenarbeit zwischen Forschung und Praxis ist hier weniger eng, der konzeptionelle Rahmen dadurch weiter gesteckt (Gräsel & Parchmann, 2004). Als Modellprogramm konzipiert, wird in SINUS die Verbreitung eines modularartig aufgebauten Ansatzes zur Qualitätsentwicklung in den Blick genommen. Der von der Forschergruppe im CHiK-Projekt geprägte Begriff der „Symbiotischen Implementationsstrategie“ soll die Vereinigung und Zusammenarbeit der Akteure im Prozess der Umsetzung von Unterrichtskonzeptionen im Chemieunterricht verdeutlichen. Die im Bereich der bottom-up-Implementation angesiedelten Strategien gehen dabei nicht von einem fertigen Produkt aus, vielmehr erfolgt eine Erprobung und anschließende Optimierung gemeinsam entwickelter Maßnahmen nach dem *trial-and-error-Prinzip*. Die Übertragung von erfahrungsbasierten Erkenntnissen bei der Implementation einer schulischen Innovation führte dazu, dass die Lehrkräfte von Anfang in die Ausgestaltung und Umsetzung der CHiK-Konzeption involviert waren. Auf der Basis von Grundideen zum Chemieunterricht entwickelten sowohl die Lehrkräfte als auch die Fachdidaktiker gemeinsam Materialien und Unterrichtsreihen, die unter den Bedingungen des jeweiligen Klassenverbandes erprobt wurden. Die Symbiose von Fachdidaktik und Erziehungswissenschaft wurde für das CHiK-Projekt durch die Bildung von Lerngemeinschaften realisiert (Abschnitt 3.2.3). Diese aus mindestens zwei Lehrkräften pro Schule bestehenden Lerngemeinschaften wurden um eine/einen Fachdidaktiker/in ergänzt und als so genanntes ‚Set‘ definiert, wobei in einem solchen Set die Lehrkräfte mehrerer Schulen nach dem Kriterium der regionalen Nähe zusammenarbeiteten. Da es sich im Projekt CHiK um die Einführung eines curricularen Ansatzes für den Chemieunterricht handelt, kooperierte jedes Set mit einer Person aus dem jeweiligen Bildungsministerium oder der Bezirksregierung. Dadurch wurde eine Kohärenz zwischen der Schulebene und der Systemebene der Mini-

sterien hergestellt, die das Projekt im Gespräch der Fachöffentlichkeit hielt und die Übereinstimmung mit aktuellen und künftigen Lehrplänen gewährleistete. In den alle zwei Monate stattfindenden Treffen wurden zunächst Materialien, Aufgaben und Unterrichtsentwürfe unter der Mithilfe der Fachdidaktiker erarbeitet, die nach dem unterrichtlichen Ausprobieren diskutiert und optimiert wurden. Die im Fokus dieses Abschnitts stehende Implementationsstrategie wurde durch ein Multiplikatorensystem umgesetzt. Ein engagiertes Mitglied der ersten Setgruppe übernahm die Koordination und Betreuung eines neu gegründeten Sets, das durch den Rückzug der Fachdidaktiker an Eigenständigkeit zunahm. Diese als Multiplikator wirkende Lehrkraft trug die Erfahrungen und Erkenntnisse der ersten Sets in die neu gegründete Runde und entwickelte somit die bis dahin entstandenen Ideen und ausprobierten Unterrichtsreihen weiter. Gelang der Rückzug der Fachdidaktiker und Fachdidaktikerinnen, trugen die Sets automatisch mehr Verantwortung, wodurch eine dauerhafte Implementation erfolgte: „...die stärkere Verantwortung der Personen aus den Ländern sowie die damit einhergehende Autonomie der Sets stellten damit eine Basis für eine dauerhafte Verankerung von *Chemie im Kontext* in der Praxis dar (Hervorhebungen im Original, Fußangel et al., 2008, S. 64). Hinsichtlich der nachhaltigen Verbreitung des Unterrichtskonzepts ist der Blick auf die Wirkungen der zweiten Sets mit erhöhter Eigenständigkeit geeignet, da hier bereits die Loslösung von der Projektgruppe stattfand und die Umsetzungshilfen der wissenschaftlichen Mitarbeiter stark zurückgenommen wurden. Fussangel et al. (2008) halten für symbiotische Implementationsstrategien eindeutig fest, dass eine Art ‚Kontrolle‘ bezüglich der Umsetzung nicht beabsichtigt ist, wodurch sich das Produkt des Chemieunterrichts variabel gestaltet. Die Autoren messen den Grad des Erfolgs der Strategie durch einen Vergleich der Lerngemeinschaften mit relativ hoher Projektgruppenanbindung zum Projektbeginn (Set 1) und relativ geringer Projektgruppenbindung in den Multiplikatorensätzen (Set 2). Der Abstand zwischen den Sets beträgt ein Schuljahr, d.h. wenn die Lehrkräfte des Set 1 im Schuljahr 2002/2003 mit der Arbeit begannen, dann erfolgte die Arbeit der Sets 2 mit den im Laufe eines Schuljahres ausgebildeten Multiplikatoren zu Beginn des Schuljahres 2003/2004. Die von Fußangel et al. (2008) ausgewerteten Daten beurteilen den Erfolg der Implementationsstrategie anhand der Befragungen zum Ende des Schuljahres 2003/2004³⁶. Aus den Ergebnissen der Lehrkraftbefragungen schließen die Autoren, dass die Dissemination

³⁶ Die ersten Sets (Set 1) arbeiteten damit bereits seit zwei Schuljahren im Projekt, die anschließend gebildeten Sets 2 erst seit einem Schuljahr.

über Multiplikatoren erfolgreich war, da der Grad der Betreuung der Sets durch die Fachdidaktiker keinen Einfluss auf die Umsetzung der CHiK-Konzeption zeigte. Die anfangs von der Projektgruppe betreuten Sets nehmen die intendierten Unterrichtsveränderungen in gleichem Ausmaß wahr, wie die Sets 2, die von den Erfahrungen und dem Wissen der Multiplikatoren profitierten. Auch hinsichtlich des von den Lehrkräften befürchteten Kontrollverlustes im Unterrichtsgeschehen bestätigen die Ergebnisse das CHiK-Konzept: „Auch in einem kontextorientierten Unterricht haben die Lehrerinnen und Lehrer nicht das Gefühl, die Kontrolle über das Unterrichtsgeschehen zu verlieren.“

(Fussangel et al., 2008, S. 71). Die von den Lehrkräften empfundenen Unterrichtsveränderungen zeigten sich in der Wahrnehmung der Schülerinnen und Schüler nach zwei Schuljahren jedoch nur bedingt. Der mit dem Konzept intendierte Rückgang von fachsystematischem Wissen zugunsten des Anwendungswissens und des selbstgesteuerten Lernens tritt zwar ein, die Lernmotivation entwickelt sich jedoch rückläufig und die Interessenentwicklung der Schülerinnen und Schüler zeigt keine Effekte. Das Autorenteam sieht das Ziel einer erfolgreichen Implementation der CHiK-Konzeption über Multiplikatoren nach insgesamt fünf Jahren als erreicht an und führt dieses Ergebnis auf das von Anfang an hohe Maß an Autonomie zurück. „Erfahrene Lehrkräfte, die neue Lehrerinnen und Lehrer in die CHiK-Arbeit einführen und sie begleiten, können in gleicher Weise die zentralen Grundideen von Chemie im Kontext verbreiten, wie dies die Fachdidaktikerinnen und Fachdidaktiker zu Beginn des Projektes getan haben.“ (Fussangel et al., 2008, S. 77). Weitere Gründe für den Erfolg werden in der inhaltlichen Nähe des in den Lerngemeinschaften kooperativ entwickelten Wissens gesehen. Für die neu entwickelten Materialien und Aufgaben ist der hohe Bezug zur alltäglichen Praxis kennzeichnend, da die Lehrkräfte selbst die Unterrichtsreihen entwickelten. Aus weiteren Teilstudien kommt die Forschergruppe zu Ergebnissen, die für die Konzeption weiterer Schulinnovationsprojekte wichtige Hinweise liefern: hinsichtlich eines Anreizes zur Projektteilnahme wird die Gewährung von Ausgleichsstunden nahezu ausgeschlossen. Schellenbach-Zell et al. (2008) räumen ein, dass Ausgleichsstunden neben der Belohnung für das Engagement auch als Druck von außen interpretiert werden, der zur engagierten Teilnahme verpflichtet, wenngleich sich in ihren Untersuchungen zur Teilnahmemotivation am Projekt keine eindeutigen Erkenntnisse zeigen. Mit Blick auf die Arbeitsbelastung von Lehrkräften bieten Ausgleichsstunden jedoch die Möglichkeit, sich langfristig zu engagieren. In der Zusammenarbeit der Lehrkräfte kommen die Autoren zu dem Ergebnis, dass innerhalb der

Fachgruppe einer Schule ein intensiverer schülerbezogener Austausch im Kollegium besteht, der verständlicherweise auf die Kenntnis des Kollegiums über die betreffenden Personen zurückzuführen ist. In den schulübergreifenden Sets erfolgen verstärkt der (fachliche) Austausch und die Kooperation zur Arbeitsorganisation. Für die Autoren stellen diese Kooperationsformen die wünschenswerte intensivere Form der Zusammenarbeit dar. Sie resümieren, „...dass sich eine Steigerung der Fachgruppenkooperation durch ein Projekt „von außen“ als schwierig erweist, wenn die Lehrkräfte zugleich in ein außerschulisches Setting integriert sind, in dem sie einer intensiven Zusammenarbeit mit Kolleginnen und Kollegen nachgehen können.“ (Schellenbach-Zell et al., 2008, S. 105). Für zukünftige Projekte scheint es daher wichtig, die schulinterne Zusammenarbeit stärker in den Implementationskonzeptionen zu berücksichtigen.

Auch für das Modellprogramm SINUS kann Ostermeier (2004) zeigen, dass die Arbeit in den Schulnetzwerken von den Lehrkräften positiv beurteilt wird. Die im Programm arbeitenden Lehrkräfte schätzen die Qualität der Zusammenarbeit, die Zufriedenheit mit der Arbeit im Programm und die wahrgenommenen Entwicklungen deutlich positiver ein als noch zu Beginn der Netzwerkarbeit³⁷. Problematisch zeigten sich in der Untersuchung die neu in die Programmarbeit eingestiegenen Lehrkräfte. Sie hatten Schwierigkeiten, einen Anschluss in Form eines Kooperationspartners auf der Netzwerkebene zu finden. Ebenso problematisierte Ostermeier (2004) die Bedeutung der betreuenden Setzkoordinatoren. Die bis dahin wenig untersuchte Rolle der Koordinierenden in einem Schulnetzwerk erwies sich als äußerst bedeutend. Bereits in der 2002 veröffentlichten Konzeption zur Dissemination des Modellprogramms spricht das Autorenteam des IPN die Wichtigkeit der den Setzkoordinatorinnen und Setzkoordinatoren innehabenden Kompetenzen an, wie z.B. Kenntnis über die Module, die Programmkonzeption, Qualitätssicherung und Unterstützungssysteme sowie Moderationstechniken. Anders als im CHiK-Projekt ging es in SINUS darum, thematisch konzipierte Module zur (mathematischen) Unterrichtsentwicklung zu bearbeiten und damit einen Qualitätsentwicklungsprozess nachhaltig zu etablieren. Als Modellversuchsprogramm angelegt, sollte sich die Konzeption wellenartig verbreiten (Abbildung 10), indem eine programmerfahrene Schule ihre Erfahrungen und ihr neues fachliches, methodisches und organisatorisches Wissen an Schulen weitergibt, die mit der Programmarbeit erst beginnen.

³⁷ Zur inhaltlichen Ausgestaltung des SINUS-Programms Abschnitt 5.1.2.

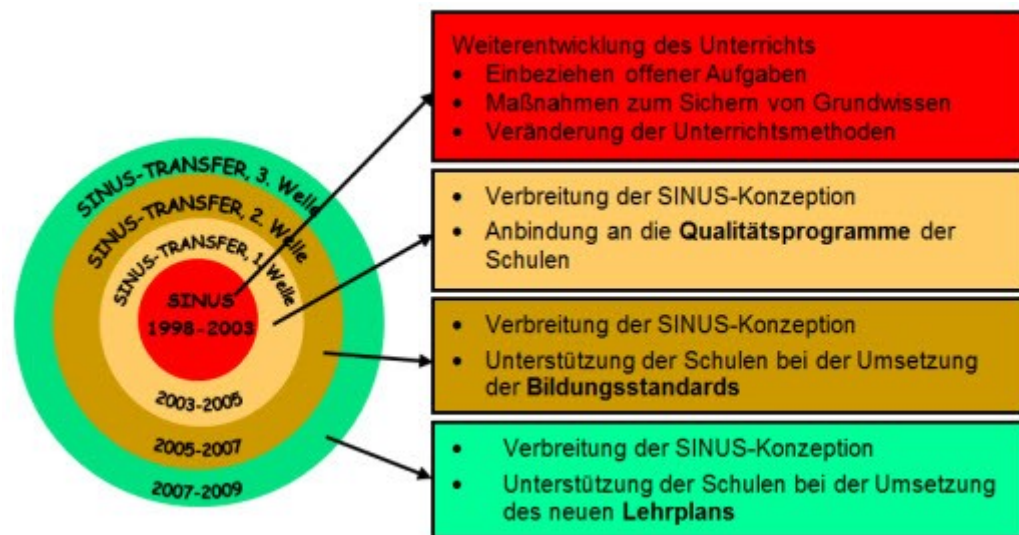


Abbildung 10. Verbreitung der SINUS-Konzeption über die Zeit, Quelle: <http://mathematik.bildung-rp.de/sekundarstufe-i/thema/sinus-und-sinus-transfer/sinus-transfer-3-2007-2009.html> [Zugriff am 23.11.2013].

Konkrete Befunde zur Implementation eines komplexen Qualitätsentwicklungsprogramms lagen aus dem deutschsprachigen Raum zum Zeitpunkt der Erstellung des Disseminationskonzeptes noch nicht vor. Doch die in Deutschland etablierten Strukturen und Ausstattungen im Schulsystem wiesen in eine pragmatische Richtung, in der konzeptionell bedacht werden musste, dass die Betreuungsressourcen des Modellprogramms für eine Verbreitung des Ansatzes nicht zur Verfügung stehen können. „Der Ansatz der kooperativen unterrichtsbezogenen Qualitätsentwicklung muss sich nach wenigen Jahren an den Schulen gewissermaßen selbst tragen“ (Prenzel et al., 2002, S. 15). So wurde die erste Disseminationswelle auf drei Jahre und für maximal 180 Disseminationssets festgelegt und in einer Verknüpfung aus den Referenzschulen des Modellversuchsprogramms mit der Ausbildung neuer Setkoordinatorinnen und –koordinatoren gesehen. Die Verantwortung für die Verbreitung des Ansatzes sollte auf die Länder übertragen werden. Im Sinne einer kohärenten Politik wurden Gelder für die Einrichtung von Koordinationsstellen und Fortbildungen bereitgestellt.

In der Evaluation der ersten Welle des SINUS-Transfer-Programms kommt klar zum Ausdruck, dass die Koordinatorinnen und Koordinatoren, die bereits im Modellversuchsprogramm aktiv waren, nunmehr vollends vom Ansatz überzeugt sind und diese Einstellung neuen Schulen vermitteln können. Dennoch halten Stadler et al. (2005, S. 18): „Die bisherigen Erfahrungen im SINUS-Transfer-Programm deuten darauf hin, dass die Verbreitung des Ansatzes zur Unterrichtsentwicklung kaum über ein Ausstrahlen von vor-

bildhaften Schulen auf neue Schulen vonstatten geht, sondern dass es sich dabei um einen komplexen kommunikativen Prozess handelt, in dem das Handeln vieler Akteure ineinander greift.“. Inwiefern der Qualitätsentwicklungsansatz auf die Bedingungen der einzelnen Schulen trifft, untersuchten Ostermeier et al. (2004). Während sich die am Programm teilnehmenden Schulen im Bereich der Naturwissenschaften in den Schülerleistungen nicht von den im Rahmen des PISA-Programms getesteten Schulen zu unterscheiden, zeigte sich im Bereich der Mathematik, dass die Schülerleistungen der Schulen mit mehreren Bildungsgängen und der integrierten Gesamtschulen bedeutend besser waren als eine vergleichbare Schülerschaft in PISA. „Dies lässt vermuten, dass sich in diesem Fällen Schulen für das Programm angemeldet bzw. gewählt wurden, die von vornherein eine gewisse Affinität zum mathematischen Bereich aufwiesen.“ (Ostermeier et al., 2004, S. 233-234). Andererseits liegen die Leistungen der Gymnasiasten in SINUS signifikant unterhalb der PISA-Vergleichsgruppe. Die Autoren schlussfolgern, dass sich Schulen mit explizitem Förderbedarf in diesem Bereich zum SINUS-Programm anmeldeten. Mit der größeren Verbreitung des Programms in den folgenden Wellen bildeten die Länder Schwerpunkte in der Entwicklungsarbeit und passten den Unterrichtsentwicklungsansatz jeweils bundeslandspezifisch an die vorherrschenden Bedingungen an. Das Bundesland Hessen nutzte die SINUS-Konzeption und deren erfolgreiche Implementation auch in der Arbeit zur Implementation der Bildungsstandards und konzipierte eine Fortbildungsreihe zum kompetenzorientierten Unterrichten (Abschnitt 5.1.2, Projektbeschreibung Hessen).

3.3.3 Studien aus der Begleitforschung zur Implementation der Bildungsstandards

Die in Deutschland aufgelegten Konzepte zur Veränderung der unterrichtlichen Praxis gaben auch für den weitreichenden Paradigmenwechsel zur Outputorientierung Impulse für die Begleitforschung zur Implementation der Bildungsstandards und der damit implizierten Innovationen (Abschnitt 5.1). Aus dem zunächst auf zwei Jahre angelegten Projekt zur „Implementation der Bildungsstandards in Berliner und Brandenburger Schulen“ berichten Pöhlmann et al. (2014) den Grad der Auseinandersetzung mit den Bildungsstandards nach einem Schuljahr. Auf der Grundlage des CBAM-Modells erfassten sie mit Hilfe eines auf die Innovation „Bildungsstandards“ adaptierten Fragebogens die kognitive Auseinandersetzung von Mathematiklehrkräften in einem Experimentaldesign. Dabei wurden Lehrkräfte erfasst, die eine Prozessbegleitung bei der Implementation der Bildungsstandards erhielten (Projektschulen) und gegen eine Kontrollgruppe ohne Unterstüt-

zung abgebildet³⁸. Im Grad der Auseinandersetzung mit der Innovation „Bildungsstandards“ unterschieden sich die beiden Gruppen signifikant voneinander. Die Lehrkräfte mit einer begleitenden Unterstützung konnten innerhalb eines Schuljahres ihre Auseinandersetzung mit den Bildungsstandards dahingehend verstärken, dass sie sich von einem selbstbezogenen Fokus entfernten und ihre Gedanken auf die Schülerinnen und Schüler richteten. So konnte die Unsicherheit bei der Konfrontation mit den Bildungsstandards durch Information und Unterstützung in Gedanken um die Wirksamkeit und Effizienz der Reformmaßnahme umgewandelt werden. „Diese Veränderung im Profil bedeutet eine tatsächliche Weiterentwicklung bei der Auseinandersetzung mit den Bildungsstandards, da im SoC-Modell die einzelnen Stufen als Prozess-Stadien formuliert werden, die nacheinander durchlaufen werden.“ (Pöhlmann et al., 2014, Abschnitt 3.1.3). Hinsichtlich der Anwendung im Unterricht zeigt sich ein ähnliches Bild. Bedingt durch einen fortlaufenden fachdidaktischen Input wendeten die Lehrkräfte in den Projektschulen die Bildungsstandards auch häufiger an, setzten sie rezeptbuchartig im Unterricht ein und kooperierten häufiger miteinander. Inwiefern eine verstärkte fachdidaktische Kompetenzorientierung erfolgte, wurde mit einem Fragebogen zur Häufigkeit von kompetenzorientierten Schülertätigkeiten gemessen. Es zeigte sich, dass die Lehrkräfte in den Projektschulen mit Ausnahme der mathematischen Kompetenzen „Argumentieren“ und „Technisch Arbeiten“ häufiger Schülertätigkeiten initiieren, die einen kumulativen Kompetenzaufbau im Sinne der Bildungsstandards fördern. Die Autoren interpretieren daraus vorsichtig, dass sich das Handeln der Lehrkräfte im Sinne der Bildungsstandards verändert hat. Einschränkend wird jedoch festgehalten, dass die Fragebögen nur die Häufigkeit des Einsatzes von kompetenzorientierten Schülertätigkeiten abfragen, womit die Frage einer geänderten Lernkultur nicht beantwortet werden kann. „Dazu würden beispielsweise die Förderung individueller Lernprozesse, kognitiv anregende Unterrichtsgestaltung, intelligentes Üben, ein veränderter Umgang mit Fehlern und entdeckendes Lernen gehören.“ (Pöhlmann et al., 2014). Bestätigt wird diese vorsichtige Interpretation durch die sich nicht veränderten Schülerleistungen. Da die Schülerleistungen durch kompetenzorientierte Aufgaben gemessen wurden, war hier eine Veränderung zugunsten der Projektschulen erwartet worden, die sich jedoch nicht einstellte. Hall und Hord (2006, S. 258) schreiben dazu: „*There will be no change in outcomes until new practices are implemented.*“ (Hervorhebungen im Original).

³⁸ Eine ausführliche Beschreibung des Projektes erfolgt in Abschnitt 5.1.

Im qualitativen Teilprojekt untersuchten Zeitler, Heller und Asbrand (2012) in einer rekonstruktiven Studie zur Implementation der Bildungsstandards, wie Lehrkräfte die Reformmaßnahme der Bildungsstandards in der Schule umsetzen. Dazu wurden unterschiedliche Fachkonferenzen aus verschiedenen Schulformen mit dem Ziel in den Blick genommen, ihre handlungsleitenden Orientierungen bei der Umsetzung der Reformidee zu rekonstruieren. Die Autorinnen fokussierten in ihren Gruppendiskussionen die Frage, wie Lehrkräfte mit der Innovation Bildungsstandards umgehen, sie verstehen und umsetzen. Für das Verständnis der Bildungsstandards können die Autorinnen zwei unterschiedliche Interpretationen von Standards ausmachen. Die Gruppe der Lehrkräfte, die Standards als eine Vorgabe betrachtet, fasst sie als vorgeschriebene und abzuarbeitende Inhalte auf, die analog zu den Lehrplänen als ein Input in das System verstanden werden. „Für diese Lehrkräfte sind Bildungsstandards keine Anregung zur Weiterentwicklung der Unterrichtspraxis, sondern die Standards werden vorrangig im Kontext ihrer Überprüfung betrachtet.“ (Zeitler et al., 2012, S. 214). Eine andere Gruppe versteht die Bildungsstandards als Aufforderung zu stärker autonomer Gestaltung von Lehr-Lern-Prozessen als bisher. Diese Lehrkräfte sehen den Gehalt der Bildungsstandards darin, dass keine konkreten Anweisungen gegeben werden und sie ihre Unterrichtspraxis selbstbestimmt gestalten können.

Im Umgang mit den Standards fanden die Autorinnen in ihrem Sample drei verschiedene Formen. Von den autonom orientierten Lehrkräften, die sich in ihrer Gestaltungsfreiheit bestätigt sehen, unterscheiden sich die heteronom orientierten Lehrkräfte, die ihre Handlungspraxis durch Vorgaben normiert sehen und sich dadurch stärker fremdbestimmt verstehen. Diese Gruppe von Lehrkräften versteht die Bildungsstandards jedoch als Autonomieaufforderung. Eine letzte Form des Umgangs mit den Bildungsstandards bildet die Gruppe der heteronom, auf habitueller Ebene geprägt orientierten Lehrkräfte, „...die Bildungsstandards als unmittelbar umzusetzende Vorgabe auffassen und somit eine Bestätigung ihrer Erwartungen an Normen erfahren.“ (Zeitler et al., 2012, S. 215). Bezogen auf den Unterricht gehen Lehrkräfte ganz unterschiedlich mit den anwendungsbezogenen Aufgaben um, die die Kompetenzen illustrieren sollen. Die mit der Reformidee implizierte Veränderung der Unterrichtspraxis hin zur Kompetenzorientierung versendet teilweise auf der Ebene der Einzelschule, da die impulsgebenden kompetenzorientierten Aufgaben für diese Lehrkräfte keinen Anlass darstellen, ihren Unterricht weiterzuentwickeln bzw. eine Änderung der Lernkultur herbeizuführen. Gründe für die Abwehrhaltung sehen die

Lehrkräfte in dem Unvermögen der Schülerinnen und Schüler, mit den anwendungsbezogenen Aufgaben umzugehen. Die in dieser Gruppe vorherrschende instruktivistisch gedachte Unterrichtspraxis führt dazu, dass die kompetenzorientierten Aufgaben, wenn gleich attraktiv, lediglich als Input verstanden werden, womit eine oberflächliche Zustimmung zur Reform „Bildungsstandards“ erfolgt. Andere Gruppen gehen konstruktiv mit den gewonnenen Freiheiten in der Unterrichtsgestaltung um und erkennen eine Anschlussfähigkeit in den von der Bildungspolitik bereitgestellten Impulsen an ihre eigene Unterrichtspraxis. Ausgehend von einem konstruktivistischen Unterrichtsverständnis reflektieren diese Lehrkräfte Möglichkeiten, „...den kumulativen Kompetenzerwerb der Schülerinnen und Schüler im Unterricht gezielt zu fördern, und können das Konzept der Kompetenzorientierung in eine entsprechende, habituell verankerte Lernkultur integrieren.“ (Zeitler et al., 2012, S. 234). Für das weitere bildungspolitische Vorgehen im Implementationsprozess der Bildungsstandards empfehlen die Autorinnen, die Heterogenität in der Lehrerschaft zu beachten. Implementationsmaßnahmen müssen den Habitus und das Vorverständnis der Lehrkräfte beachten. Im Wesentlichen bestätigen und unterstützen die Autorinnen ein Vorgehen analog zum SINUS-Konzept, das den autonom orientierten Lehrkräften eine Partizipationsmöglichkeit und Anerkennung bietet und den heteronom orientierten Lehrkräften Sicherheit und Klarheit im Umgang mit der empfundenen Herausforderung verschafft. Für die problematische Gruppe der heteronom orientierten Lehrkräfte, die Bildungsstandards als eine „Weisung von oben“ verstehen, scheinen längerfristige individuelle Prozessbegleitung und entsprechende Fortbildungsangebote Erfolg versprechend zu sein, die Lehrkräfte an ein Autonomieverständnis heranzuführen, wodurch die Bildungsstandards auch einen Innovationsgehalt erhalten.

Folgt man den Ausführungen von Zeitler et al. (2012) erscheint es sinnvoll, vor der Konzipierung einer Fortbildungsmaßnahme eine „Bestandsaufnahme“ zur Orientierung der Lehrkräfte durchzuführen. Aus ökonomischen Gründen bieten sich Fragebögen zu den Einstellungen gegenüber der Innovation an. Auch die im quantitativen Teilprojekt eingesetzten Instrumente des „Stages of Concern“ oder des „Level of Use“ Fragebogens bieten die Möglichkeit, Lehrkräfte zu kategorisieren und daraufhin eine geeignete Unterstützung oder Beratung zu konzipieren. Zeitler et al. (2012) und Pöhlmann et al. (2014) konnten in Übereinstimmung mit den bestehenden Theorien die Bedeutung der Faktoren Einstellungen, Orientierungen und Kognitionen von Lehrkräften empirisch hervorheben. Dass eine Unterstützung bei der Umsetzung bildungspolitischer Innovationen dazu führt, sich mit

der Innovation zu beschäftigen und die fachdidaktischen Inputs im Unterricht auszuprobieren, zeigt die Studie von Pöhlmann et al. (2014). Inwiefern sich daraus jedoch eine nachhaltige Veränderung der Lernkultur ergibt, vermag diese Publikation nicht vorherzusagen. Hier scheinen die Ergebnisse des qualitativen Teilprojekts schon eher einen Erklärungsgehalt zu liefern. In Abhängigkeit von der handlungsleitenden Orientierung der Lehrkräfte werden die Bildungsstandards unterschiedlich verstanden. Die Ergebnisse des quantitativen Teilprojektes könnten hier weitere Erkenntnisse liefern.

4 Fragestellung

Die umfangreiche Literatur zur Implementation von Bildungsinnovationen wurde in Kapitel 3 untersucht. Es zeigte sich hierbei, dass neben der Fähigkeit zum lebenslangen Lernen die Kooperation in professionellen Lerngemeinschaften ein vielversprechender Ansatz zur Erprobung und Umsetzung neuer Konzepte im Bereich der Unterrichtsentwicklung darstellt. Eine Adaption der in Abschnitt 3.3.2 vorgestellten Projektkonzeptionen von SINUS und CHiK erscheint für die Einführung eines kompetenzorientierten Unterrichts daher sinnvoll. Die erste übergeordnete Forschungsfrage wird daher folgendermaßen formuliert:

I. Welche Effekte werden durch eine gezielte Professionalisierungsmaßnahme zur Implementation der Bildungsstandards bei den Lehrkräften beobachtet?

Den in Abschnitt 3.2.1 folgenden Ausführungen unterliegen die Hypothesen, dass Lehrkräfte ihre Einstellungen und darauf aufbauend ihr Unterrichtsverhalten ändern können. Gleichzeitig können die Hypothesen jedoch auch andersherum formuliert werden und untersuchen, ob Lehrkräfte zunächst die neuen Lehrmethoden ausprobieren und danach erst ihre Einstellungen ändern.

Die in der angestrebten Professionalisierungsmaßnahme vorausgesetzte Teamarbeit kann einen Hinweis auf die Effizienz der Arbeit in den Schulen geben. Wie die Erkenntnisse in Abschnitt 3.2.3 zeigen, probieren Lehrkräfte neue Lehrmethoden schneller aus, wenn die Isoliertheit und das Einzelkämpferkultur überwunden werden. Die Hypothese hinsichtlich einer erhöhten Zusammenarbeit innerhalb der Professionalisierungsmaßnahme kann durch den Grad der Kooperation gemessen werden. Aufgrund der bisherigen Forschungslage wird eine verstärkte Kooperation in den Lehrerteams erwartet.

Damit verbunden wird ebenfalls eine inhaltlich intensivere Arbeit erwartet, die über die Ziele eines qualitativ hochwertigen Unterrichts gemessen werden soll.

Fullan (2007) zeigt in seinen Ausführungen, dass eine Änderung der Praxis mindestens die drei Aspekte

- Einsatz neuer Materialien
- Einsatz und Erprobung neuer Lehrmethoden
- Änderung der Einstellungen

umfasst. Dabei definiert er den Erfolg einer Innovation: „Success is not just about being right; it is about engaging diverse individuals and groups who likely have many different versions about what is right and wrong.“ (Fullan, 2007, S. 40). Die Implementation der Bildungsstandards wird im Folgenden konzeptionell durch eine Professionalisierungsmaßnahme gefördert, in der die Verwendung neuer Materialien (Aufgaben) und das Ausprobieren neuer Lehrmethoden im Mittelpunkt stehen. Die mit der Umstellung auf Kompetenzen verbundene neue Lehrausrichtung führt zu der Erwartung, dass Lehrkräfte kompetenzorientierte Tätigkeiten in ihrem Unterricht häufiger initiieren und diese über einen längeren Zeitraum auch verstetigen. Die Zieldefinition eines standardorientierten Unterrichts wurde im Kapitel 2 beschrieben. Daraus ergibt sich die zweite übergeordnete Forschungsfrage:

II. Welche Auswirkungen eines kompetenzorientierten Unterrichts können auf Seiten der Schülerinnen und Schüler beobachtet werden?

Erwartet wird, dass Schülerinnen und Schüler häufiger kompetenzorientierte mathematische Tätigkeiten in ihrem Unterricht wahrnehmen. Die erwartete Veränderung in der Wahrnehmung der Unterrichtsdurchführung hängt eng mit der Lehrperson zusammen. So wird weiterhin untersucht, wie Schülerinnen und Schüler ihre Lehrkräfte hinsichtlich ihres professionellen Auftretens bewerten. Der Bezug zu den Einstellungen der Lehrkräfte liegt hier nahe. Ebenso werden die Mathematikleistungen untersucht, die ursprünglich zur Einführung von Bildungsstandards führten. Ein qualitativ hochwertiger Mathematikunterricht soll sich letztendlich auch in den Leistungen der Schülerinnen und Schüler niederschlagen. Erwartet wird daher eine Steigerung des Leistungsniveaus über die Zeit.

Die bisherigen Ausführungen legen für eine politisch intendierte Paradigmenwende im Bildungswesen nahe, die Verhaltensweisen der an Schule beteiligten Personen, ihre Einstellungen, ihre Praktiken und Beurteilungen dahingehend zu verändern, den neuen

Blickweisen gerecht zu werden. Den wahrscheinlichsten Eintritt eines veränderten Lehrhabitus sehen Oelkers und Reusser (2008) in langfristigen Maßnahmen auf verschiedenen Ebenen des Schulsystems zur Sensibilisierung und Professionalisierung des pädagogischen Personals.

Mit Bezug auf das Zusammenspiel von Bildungsstandards und Bildungsevaluation weist Drieschner (2010) darauf hin, dass es sich Lehrkräfte nicht mehr leisten können, an den Bildungsstandards vorbei zu unterrichten. Die durch stetige Vergleichsstudien und Evaluationen sichtbar werdenden Entwicklungserfolge der Schulen, lassen die Bedeutung von Handlungsalternativen bei der Einführung der Bildungsstandards steigen, d.h. wenn den Schulen ein Konzept bekannt ist, nach dem es ihnen gelingen kann, den Unterricht hinsichtlich der auszubildenden Kompetenzen am Ende der Schulzeit umzustellen, kann eine Implementation mittelfristig erfolgreich durchgeführt werden. Das Wissen über förderliche Faktoren, unterschiedliche Rahmenbedingungen und den Umgang mit einem heterogenen Schülerklientel in Implementationsprozessen würde die Bundesländer bei der Konzeption geeigneter Fortbildungen oder Implementationsstrategien unterstützen. Die bisherigen Aktivitäten in den Bundesländern weisen einzelne Angebote auf, lassen eine Systematik implementierender Maßnahmen jedoch vermissen (Oelkers & Reusser, 2008). Die Generierung von Wissen über Verhaltensänderungen bei der Einführung einer bildungspolitischen Innovation ist daher eines der zentralen Anliegen aller mit der Implementationsaufgabe betrauten Personenkreise. Am IQB etablierte sich neben der Hauptaufgabe der Entwicklung von Kompetenzmodellen der Anspruch, die Implementation der Bildungsstandards begleitend zu beforschen. Wenngleich die Aufgabe der Einführung und Umsetzung dieser Innovation bei den Bundesländern liegt, so versteht sich das IQB als wissenschaftliche Einrichtung der Länder auch als die Institution, die für den Prozess der Implementation wissenschaftliche Expertise aufbaut und bereitstellt, die neben vorliegenden Fortbildungsmaterialien auch Forschungsergebnisse und Handlungsempfehlungen umfasst. Im empirischen Teil dieser Arbeit wird zunächst die Begleitforschung zur Implementation der Bildungsstandards am IQB konzeptionell vorgestellt, bevor die Datensätze beschrieben und im Hinblick auf die Beantwortung der Forschungsfragen analysiert werden. Eine Einordnung der Ergebnisse und deren Diskussion im Kontext bildungspolitischer Reformmaßnahmen erfolgt in Kapitel 7.

Empirischer Teil

5 Die Begleitforschung zur Implementation der Bildungsstandards

Eine im Jahr 2006 durchgeführte Recherche zu den Implementationsaktivitäten in den Bundesländern ließ keine konkreten Projekte erkennen, in denen der Umgang mit den Bildungsstandards thematisiert wurde, wenngleich Schleswig-Holstein mit den Regionalen Qualitätsforen und Hessen mit der Konkretisierung des SINUS-Programms erste Schritte in diese Richtung unternahmen. So wurde in den Jahren 2007 bis 2009 am IQB ein Projekt mit dem Ziel durchgeführt, ein Konzept der Implementation der Bildungsstandards zu entwickeln, das nach einer Erprobung und Evaluation Modellcharakter haben könnte. Inhaltlich sollten die Möglichkeiten eines standard- bzw. kompetenzorientierten Unterrichts aufgezeigt und weiterentwickelt werden. Das Implementationskonzept sollte durch

1. eine fachspezifische Unterrichtsentwicklung zu
2. einem output-orientierten Unterricht und
3. einer Verinnerlichung der Kompetenzorientierung bei der Unterrichtsgestaltung

führen. Die fortgeschrittenen Arbeiten im Fach Mathematik Sekundarstufe I gaben den Ausschlag dafür, das Konzept inhaltlich für das Fach Mathematik zu erproben und im weiteren Verlauf zu entscheiden, ob eine Ausweitung auf andere Fächer erfolgen soll. Konzeptionell an das SINUS-Programm angelehnt, sollten die Lehrkräfte einen Unterrichtsentwicklungsprozess beginnen, der durch eine Koordinatorin moderiert und begleitet wird. Lehrkräfte werden als Experten für Unterricht angesehen, die auf der Grundlage von Modulen eine thematisch fokussierte Weiterentwicklung des eigenen Unterrichts in der Fachgruppe betreiben sollten und dafür durch geeignete Unterstützungssysteme Anregungen erhielten. Die quantitative Begleitforschung erfolgte in diesem Zeitraum durch die Erhebung von Lehrerdaten und den Einsatz von standardbasierten Mathematiktests und Fragebögen bei den Schülerinnen und Schülern der anfänglich neunten Jahrgangsstufe. Insgesamt erfolgten fünf Datenerhebungen, die den Entwicklungsprozess hinsichtlich des Ziels eines kompetenzorientierten Unterrichts abbilden sollten. Der Projektzeitraum von insgesamt fünf Schulhalbjahren teilt sich in zwei Teilstudien. Die erste Teilstudie umfasst die ersten drei Schulhalbjahre, in denen jeweils eine Datenerhebung erfolgte. Die Schülerinnen und Schüler in dieser Teilstudie befanden sich zum Zeitpunkt der ersten Datenerhebung in der neunten Jahrgangsstufe und beendeten nach einem Zeitraum von drei Schulhalbjahren meist die Sekundarstufe I.

Die Teilstudie zwei setzte im Schuljahr 2008/2009 an und umfasste bis zum Ende des Projektes im Juli 2009 zwei weitere Messzeitpunkte. Wie in Abbildung 11 dargestellt, wurde in der zweiten Teilstudie eine gemischte Schülerstichprobe aus neunter und zehnter Jahrgangsstufe getestet und befragt. Weiterhin können die teilnehmenden Schulen hinsichtlich der stattfindenden Intervention unterschieden werden. Die Stichprobe besteht einerseits aus Projektschulen, die an der Erprobung des IQB-Implementationskonzeptes teilnehmen und andererseits aus Schulen, die am SINUS-Programm teilnehmen. Den Schulen, in denen Entwicklungsprozesse beobachtet werden sollten, wurden jeweils weitere Schulen ohne Entwicklungsprojekte gegenübergestellt (Vergleichsschulen).

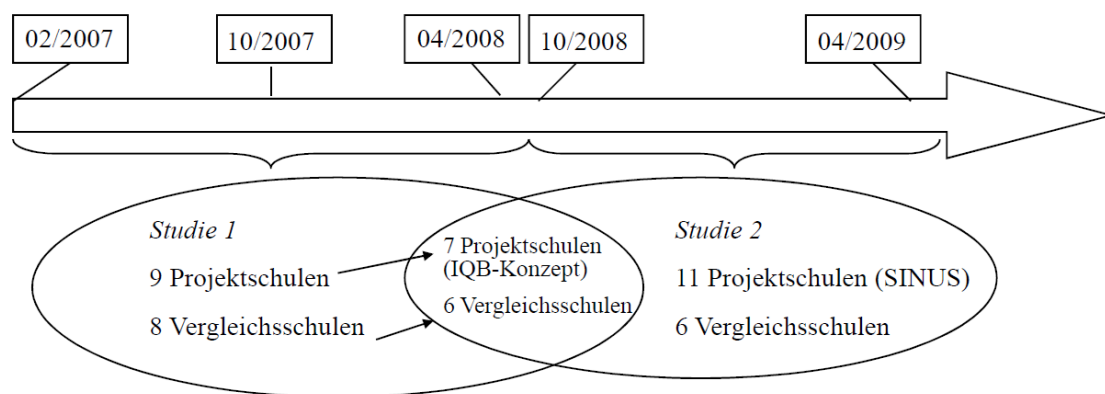


Abbildung 11. Untersuchungsdesign der quantitativ-standardisierten Begleitforschung im Zeitraum 2007-2009; Quelle: eigene Darstellung.

Hinsichtlich des Untersuchungsdesigns ist zu erkennen, dass es sich bei beiden Studien um ein Experimentaldesign mit Messwiederholung handelt. Durch die Erhebung von Daten in so genannten Kontroll- oder Vergleichsschulen können mögliche Effekte in der Veränderung der untersuchten Merkmale auf die Intervention (hier: das IQB-Implementationskonzept und die Teilnahme an SINUS) zurückgeführt werden. Die Trennung der Begleitforschung in zwei Teilstudien erfolgte aufgrund zweier Gegebenheiten:

1. die Befragung und Testung einer neuen Schülerkohorte nach drei Schulhalbjahren aufgrund der Beendigung der Sekundarstufe I der anfänglichen Schülerkohorte,
2. die Entscheidung zur Aufnahme einer weiteren Vergleichsgruppe aus dem SINUS-Programm im Bundesland Hessen zur Evaluation möglicher Effekte beider Interventionen.

Die Teilstudien wurden damit einerseits nach einem zeitlichen Gesichtspunkt und andererseits nach dem Gesichtspunkt des Stichprobenumfangs und der Interventionsart ge-

trennt. Es bleibt jedoch zu beachten, dass die Intervention der Teilstudie 1 (Erprobung eines Konzeptes zur Implementation der Bildungsstandards) über den Zeitraum der gesamten Begleitforschung mit stabiler Teilnahme der Schulen fortgeführt wurde³⁹. Die inhaltliche Ausgestaltung der beiden Interventionen erfolgt in den folgenden Abschnitten, wobei die Intention der Programme in ihrer Gesamtheit beschrieben wird und nicht auf den Untersuchungszeitraum begrenzt ist.

5.1 Das Projekt zur „Erprobung Konzeptes zur Implementation der Bildungsstandards im Fach Mathematik in Berliner und Brandenburger Schulen“ (Teilstudie 1)

Das Wissen über erfolgreiche Unterrichtsentwicklungsprogramme wie SINUS führte am IQB dazu, die Erkenntnisse als handlungsleitend bei der Konzeption eines Programms zur Implementation der Bildungsstandards zu betrachten. Das als Kooperationsprojekt zwischen dem IQB und dem Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM) entwickelte Konzept trägt die Idee, Prozesse des gemeinsamen Reflektierens über Unterricht anzuregen und zu begleiten (Zeitler et al., 2009). Das Projekt zur „Erprobung eines Implementationskonzeptes in Berliner und Brandenburger Schulen“ wurde unter Beteiligung führender Innovations- und Implementationsforscher sowie Lehrkräften und Fortbildnern vorbereitet und anhand einer kleinen Stichprobe beispielhaft im unmittelbaren Umkreis des IQB erprobt. Das geplante Konzept sollte in Abhängigkeit von den erzielten Erkenntnissen in Form eines modellhaften Beispiels auch für die Implementation der Bildungsstandards in anderen Fächern übertragen werden.

5.1.1 Projektbeschreibung

Die Ergebnisse eines Planungsworkshops am IQB in 2006 dienten als Grundlage zur Entwicklung eines Konzeptes, das sich in Anlehnung an die Fortbildungsreihe SINUS aus einzelnen Modulen zusammensetzt und von interessierten Fachkonferenzgruppen in ausgesuchten Pilotschulen bearbeitet und erprobt wird. Die Arbeit sollte in den Schulen erfolgen und durch das IQB mit Angeboten zu zentralen Fortbildungen beim Kooperationspartner im LISUM verstärkt werden. Für die Arbeit innerhalb der Schulen wurden die teilnehmenden Fachkonferenzen von einer Koordinatorin aus dem LISUM betreut, die

³⁹ D.h. mit Beginn der Teilstudie 2 verblieben die teilnehmenden Schulen aus der Teilstudie 1 im Sample und wurden um die SINUS-Schulen ergänzt.

durch Fortbildung, Moderation und Unterstützung einen Unterrichtsentwicklungsprozess in den Schulen begleitet. Die Steuerung dieses Prozesses durch die Koordinatorin sollte das Ziel verfolgen, die Lehrkräfte in den Schulen mit dem Konzept der Bildungsstandards vertraut zu machen. Damit sollte sowohl eine Reflektion der Lehrkräfte über ihren eigenen Unterricht erfolgen als auch der Anstoß zur Entwicklung von Ideen, wie ein standardbasierter Unterricht in der Praxis realisierbar wäre. Mit der Entscheidung des IQB, die teilnehmenden Lehrkräfte als Experten für ihren Unterricht anzusehen, sollte der Anteil instruktivistischer Fortbildung minimal gehalten und durch Kommunikation sowohl innerhalb der Fachkonferenzgruppe als auch zwischen den Fachkonferenzen und mit den Koordinatorinnen substituiert werden.

Die Arbeit zwischen den Schulen erfolgte im Sinne der Netzwerkbildung auf den zentral organisierten Fortbildungsveranstaltungen. Den Lehrkräften wurde damit die Möglichkeit geboten, ihre bis dahin erlangten Erfahrungen und Erkenntnisse in diesem Pilotprojekt miteinander auszutauschen und sich über verschiedene Aspekte der Schulentwicklungsarbeit zu verständigen. Diese zentralen Veranstaltungen folgten dem Muster eines meist wissenschaftlichen Inputs mit anschließendem Workshop. Während der gesamten Projektlaufzeit von fünf Schulhalbjahren wurden insgesamt drei solcher zentralen Veranstaltungen angeboten, wobei die Zielgruppe hauptsächlich die Fachgruppen der Projektschulen waren. Tabelle 3 stellt die Themenschwerpunkte dar. Die Wertschätzung und Motivierung der Fachgruppen erfolgte zu Beginn des Projektes mit einer Auftaktveranstaltung im Februar 2007, die durch eine Veranstaltung zum Themenschwerpunkt „Entlastung durch veränderte Arbeitsorganisation“ im November 2007 verstärkt wurde, zu der auch die Schulleitungen geladen wurden.

Tabelle 3. Zentrale Veranstaltungen im Projekt zur "Erprobung eines Fortbildungskonzeptes in Berliner und Brandenburger Schulen"

Zeitpunkt	Inhaltliche Beschreibung	Zielgruppe
Februar 2008	Effektive Gestaltung von Teamsitzungen der Fachkonferenzen	Fachgruppen
November 2009	Feedback-Kultur	Fachgruppen
März 2009	Mit Unterschieden rechnen – Individualisieren und Differenzieren als Thema der Fachkonferenz Mathematik	Fachgruppen

Den Fachgruppen war es durch drei zusätzliche Netzwerktreffen möglich, in den inhaltlichen Erfahrungsaustausch mit anderen Schulen zu gelangen. Die Netzwerktreffen fanden jährlich statt und begannen im Mai 2007 mit einer Klausurtagung zu den Bildungsstandards. Nach drei Schulhalbjahren fand am Ende des Schuljahres 2007/2008 im Juni 2008 ein „Markt der Möglichkeiten“ statt. Auf dieser Veranstaltung erfolgte eine erste Rückmeldung aus dem quantitativen Teilprojekt und die Schulen präsentierten ihre ersten Arbeitsergebnisse aus den Unterrichtsentwicklungsprozessen. Im Juli 2009 wurden auf dem letzten Netzwerktreffen die Gesamtergebnisse und Erfahrungen ausgetauscht, und um einen Input zum „Kompetenzorientierten Üben“ ergänzt, der die Fachgruppen zur Nachhaltigkeit ihres Wissens anregen sollte. Diese Veranstaltungen sollten die Prozessqualität und die Kooperation in den Fachgruppen fördern mit Aspekten von Schulentwicklung bzw. insbesondere der Kommunikation und Teamentwicklung. Das Angebot des Projekts sollte weniger an die einzelne Fachgruppe angepasst sein, sondern als klassische Fortbildung mit externen Referenten verstanden werden, die festgelegte Inhalte bearbeiten. Da sich das Angebot an alle Projektschulen gemeinsam richtete, bestand auf diesen Veranstaltung die Möglichkeit zum schulübergreifenden Austausch. In der Projektkonzeption ist jedoch vorgesehen, die Kommunikation zwischen den Projektschulen nicht zu forcieren. Das Kommunikationsangebot im Sinne eines Netzwerks kann einerseits inspirierend sein, andererseits aber auch zur Überfrachtung des Projekts und zeitlichen oder inhaltlichen Überforderung der Lehrkräfte führen. Der Schwerpunkt liegt auf der Arbeit in den Schulen mit der Förderung der Kooperation innerhalb der Fachgruppe und deren Einbindung in das Gesamtkollegium. Die prozessbegleitende, inhaltliche Arbeit wurde thematisch durch die Module bestimmt, die das Ergebnis des anfänglichen Planungsworkshops bildeten (Zeitler et al., 2009).

5.1.2 Modulbeschreibung

Durch die Vorreiterrolle in der Entwicklung von Kompetenzmodellen und Implementationsmaterialien im Bereich der Mathematik (Sek. I) konnte eine enge Koppelung an das Standardwerk von Blum et al. (2006) „Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen“ erfolgen, das während der gesamten Projektlaufzeit sowohl den Schulen als auch den Koordinatorinnen als Orientierung in der Weiterarbeit dienen sollte. Die aus dem Planungsworkshop generierten acht Module wurden durch diese Vorlage inhaltlich ausgestaltet. Die Module werden im Folgenden skizziert:

Mit dem Modul (1) *Bildungsstandards* sollten sich die teilnehmenden Lehrkräfte erstmals mit dem Konzept der Bildungsstandards beschäftigen und die Aspekte einer Qualitätssicherungsstrategie auf der Makroebene nachvollziehen. Die Lehrkräfte erhalten hier Informationen über das Zusammenspiel von Tests, Kompetenzmodellen, Evaluationen und die intendierte Wende im Bildungswesen. Weiterhin gilt es, den Begriff der Kompetenz zu erörtern und die für einen kompetenzorientierten Lernbegriff notwendigen Konzepte anzusprechen. Ein dritter Aspekt dieses ersten Moduls liegt in der transparenten Darstellung der Implementationsstrategie, um den Lehrkräften den Erwartungshorizont aufzuzeigen und die weiteren Arbeitsschritte zu erklären (Teilnahme an den zentralen Fortbildungsveranstaltungen, Arbeit in der Fachkonferenz usw.).

Das Modul (2) *Aufgaben* sollte die Lehrkräfte mit der Konzeption kompetenzorientierter Aufgaben vertraut machen. Durch die Analyse von bereits veröffentlichten Aufgaben zur Überprüfung der Bildungsstandards und auch anhand der Beispiele von Lernaufgaben im Begleitbuch „Bildungsstandards Mathematik: konkret“ (2006) sollten die Lehrkräfte angeregt werden, bereits vorliegende Aufgaben in ihrem eigenen Unterricht einzusetzen und zu erproben. Weiterhin sollte in der Fachkonferenz gemeinsam daran gearbeitet werden, Lehrbuchaufgaben kompetenzorientiert umzuformulieren und ein Bewusstsein für die zu erwerbenden allgemeinen mathematischen Kompetenzen geschaffen werden. Dazu gehört die Gegenüberstellung verschiedener Aufgabentypen und Differenzierung zwischen Test- und Lernaufgaben. Das Modul hatte das Ziel, die Lehrkräfte mit den Merkmalen und Grundprinzipien einer kompetenzorientierten Aufgabenentwicklung vertraut zu machen und sie zu ermutigen, eigene Aufgaben für die weitere Anwendung im Unterricht zu entwickeln.

Im Modul (3) *Entdecken und Üben* sollten die Lehrkräfte mit den förderlichen Lernumgebungen zum Aufbau mathematischer Kompetenzen vertraut gemacht werden. Der Prämisse folgend, dass der Kompetenzerwerb ein aktiver, selbstgesteuerter Prozess der Schülerinnen und Schüler ist, wurde hier über kompetenzorientierte Unterrichtseinstiege und Unterrichtsmethoden informiert. Die von den Setkoordinatorinnen verwendeten Materialien deuten darauf hin, dass die Fachkonferenzen bei der Bearbeitung des Moduls ihr Wissen über Lernstrategien auffrischten und neue Anregungen zur Schaffung kognitiv aktivierender Lernumgebungen erhielten. Konkret wurden die Lehrkräfte mit Unterrichtsmethoden wie dem „Intelligente Üben“ (Wynands, 2006), der „Projektorientierung“, dem Einsatz mathematischer Software oder auch mit Wochenplanmethoden bekannt ge-

macht. Eine Verbindung der Module 2 und 3 ergibt sich beispielsweise bei der Methode des „Entdeckenden Üben“, in der den Lehrkräften neben geeigneten Aufgaben auch deren Variation in Form von Erweiterungen oder Umformulierungen präsentiert werden.

Die Aufdeckung von Kompetenzprofilen und die weiterführende Arbeit in Form von individueller Förderung stellt den Schwerpunkt des Moduls (4) *Diagnose und Fördern* dar. Hier sollen die Lehrkräfte damit vertraut gemacht werden, wie eine kompetenzbasierte Diagnose auf der Grundlage von Kompetenzbeschreibungen erfolgt. Dazu sollen die Lehrkräfte erkennen, welches Potential in kompetenzorientierten Aufgaben steckt und wie diese richtig eingesetzt werden. Eine Sensibilisierung für die im Kompetenzmodell Mathematik verankerten Dimensionen der Leitideen, Kompetenzen und Anforderungsbereiche ist dabei unerlässlich. Anders als in der gängigen Diagnostik muss den Lehrkräften bewusst gemacht werden, dass es beim Einsatz kompetenzorientierter Aufgaben zu diagnostischen Zwecken darum geht, so genau wie möglich zu ermitteln, inwieweit die Schülerinnen und Schüler über mathematisch-kognitive Werkzeuge verfügen (Sjuts, 2006).

Im Modul (5) *Leistungsmessung* geht es in erster Linie um die Verwendung kompetenzorientierter Aufgaben und die anschließende Bewertung der gezeigten Schülerfähigkeiten im praktischen Schulalltag. Besonders die Konzeption von Klassenarbeiten auf der Grundlage kompetenzorientierter Aufgaben steht im Mittelpunkt der Arbeit. Neben der Vorstellung möglicher Analyseschemen fließen in dieses Modul auch didaktische Überlegungen zur Leistungsbewertung und Aspekte der Rückmeldung mit ein.

Mit dem Modul (6) *langfristiger Kompetenzaufbau* soll erreicht werden, dass die Lehrkräfte den Kompetenzaufbau der Schülerinnen und Schüler über mehrere Jahre im Blick behalten und das Konzept der standardbasierten Unterrichtsgestaltung nicht nur auf die Jahrgangsstufe beziehen, die für die Überprüfungstests vorgesehen ist. Durch die Verankerung von Reflexionsphasen während und nach der Bearbeitung kompetenzorientierter Aufgaben soll den Lehrkräften aufgezeigt werden, dass es für einen langfristigen Kompetenzaufbau wichtig ist, nicht nur die Ergebnisse aus individueller oder Gruppenarbeit zu vergleichen, sondern „...die eingesetzten mathematische Begriffe, Zusammenhänge und Verfahren im jeweiligen Anwendungskontext als Mathematisierungsmuster zu erkennen und explizit herauszuarbeiten.“ (Bruder, 2006, S. 137).

Im Gesamtsystem der output-orientierten Steuerung bilden Bildungsstandards den Referenzpunkt für eine Reihe von Maßnahmen. So soll im Modul (7) *Entwicklung eines schul-internen Curriculums* auf einen aktuellen Arbeitsschwerpunkt der Schulen eingegangen

werden. Lehrkräften sollen Bildungsstandards als Anstoß für die Unterrichtsentwicklung verstehen. Im Modul sollen Umsetzungsmöglichkeiten des Unterrichtsentwicklungsprozesses im Schulentwicklungsprozess erarbeitet werden. Dabei soll aufgezeigt werden, wie ein schuleigenes Curriculum gestaltet werden kann. Die Lehrkräfte werden dazu angehalten, das Curriculum als Arbeit der Fachkonferenz zu betrachten und die Ergebnisse des Abstimmungsprozesses zu dessen Inhalten als verbindlichen Arbeitsrahmen für die didaktische Arbeit in einem kompetenzorientierten Unterricht anzusehen.

Das Modul (8) *Lebensweltbezug* fokussiert auf den inhaltlichen Schwerpunkt „Authentizität“ einer kompetenzorientierten Aufgabe und soll den Lehrkräften Möglichkeiten aufzeigen, wie mathematische Probleme in einen alltäglichen Kontext eingebettet werden können. Die Lehrkräfte sollen so für die Verwendung und Konstruktion von Aufgaben motiviert werden, die einen lebensweltlichen Bezug haben, um den Schülerinnen und Schüler so ein nützliches Bild von Mathematik zu vermitteln und den Aufbau mathematischer Kompetenzen zu unterstützen.

5.1.3 Die Umsetzung der Konzeption

Wenngleich sich durch den inhaltlichen Innovationscharakter des Konzeptes bereits die Suche nach geeigneten Trägern dieses Pilotprojektes als schwierig gestaltete, konnten doch drei Koordinatorinnen gewonnen werden, die sich auf die Betreuung von neun Projektschulen verteilten. In regelmäßig stattfindenden Koordinationstreffen zwischen den Vertretern der Konzeptentwicklung (IQB) und den Koordinatorinnen wurde deutlich, dass die Rolle der Koordinatorinnen in den Schulen missverstanden wurde. Die Veranstaltungen zur Arbeit an den Modulen konnten als Fortbildungen charakterisiert werden, die den Lehrkräften eher wenig Raum für die eigene Entwicklung ließ und top-down initiierte Entwicklungen vorgab. In der Reflexion der Koordinatorinnen zum Gesamtprojekt wird das Vorgehen beschrieben als die Kombination eines Inputs „...mit einer Anleitung zum konkreten Transfer auf die eigene Unterrichtssituation“ (LISUM, 2009, S. 35).

Neben den im Hinblick auf die Projektintention obligatorisch zu bearbeitenden Modulen 1 „Bildungsstandards“ und 2 „Aufgaben“, erfolgte die Bearbeitung weiterer Module aus den Bedürfnissen und Präferenzen der Schulen heraus. Die teilnehmenden Schulen kamen je nach organisationaler Machbarkeit und kollegialem Bedürfnis sechs bis neun Mal innerhalb von fünf Schulhalbjahren zusammen. Die Projektarbeit an den Modulen erfolgte außerhalb der regulären Unterrichtszeit, wodurch sich die Dauer der Fachkonferenztreffen auf ein bis zwei Stunden beschränkte. Der Dokumentation der Koordinatorinnen folgend,

wurden in den Schulen nach Möglichkeit alle Module angesprochen, wobei die Module ineinander übergingen. Eine explizite Trennung der thematischen Module war weder gewollt noch ist eine streng inhaltliche modulare Fokussierung in der mehrdimensionalen Institution Schule realisierbar (LISUM, 2009). Die teilnehmenden Schulen versprachen sich mehrheitlich eine Verbesserung der Schülerleistung und eine verbesserte Zusammenarbeit in der Fachkonferenz vom Projekt. In der Dokumentation der Sitzungen zeigt sich, dass in den Brandenburger Schulen im Durchschnitt fünf Module pro Treffen bearbeitet wurden und stets eine Hausaufgabe zur nächsten Sitzung von den Lehrkräften erbracht werden sollte. In allen Schulen wird eine starke Führung der Koordinatorinnen sichtbar, die sich durch die Form und Methodenanwendung bei der Darbietung des Inputs zeigte.

5.1.4 Der Ablauf der Begleitforschung

In dem Gesamtprojektzeitraum von fünf Schulhalbjahren wurden fünf Testungen und Befragungen durchgeführt. Die Gesamtlaufzeit des Projektes wurde aufgrund des Studiencharakters der Begleitforschung in zwei Zeitabschnitte unterteilt⁴⁰. Der in der Studie 1 betrachtete Zeitraum umfasst die ersten drei Schulhalbjahre bzw. eine Zeitspanne von Februar 2007 bis Juli 2008. Die Studie 2 umfasst die Messzeitpunkte vier und fünf, die in der Zeit von Oktober 2008 bis April 2009 stattfanden. Während der gesamten Projektlaufzeit arbeiteten die Fachkonferenzen an ausgesuchten Modulen und trafen sich je nach Kapazität ein bis zwei Mal pro Schulhalbjahr. Die Arbeit an den Modulen erfolgte außerhalb der regulären Unterrichtszeit, wodurch sich die Dauer der Fachkonferenztreffen auf ein bis zwei Stunden beschränkte. Nach den ersten drei Schulhalbjahren (Studie 1) schieden zwei Projektschulen aus dem Projekt aus, wobei eine Schule aus Gründen der Überlastung die Projektarbeit beendete und eine andere durch den Wegfall der Schülerschaft von der Schulauflösung bedroht war. Mit Beginn des Projektes wurden acht weitere Schulen als Vergleichsschulen akquiriert, die keine Unterstützung im Unterrichtsentwicklungsprozess erhielten und nur mit den öffentlich zugänglichen Materialien arbeiten konnten. Ein Vergleich von Projekt- und Vergleichsschulen soll den Effekt des Konzeptes verdeutlichen. In den acht zusätzlich akquirierten Schulen wurden die gleichen Testinstrumente eingesetzt (Abschnitt 5.3). Abbildung 12 zeigt, dass nach drei Datenerhebungen auch in dieser Gruppe zwei Schulen eine weitere Mitarbeit verweigerten bzw. nicht

⁴⁰ Studie 1 umfasst die ersten drei der insgesamt fünf Messzeitpunkte. Für Studie 2 wurde die Stichprobe um eine weitere Vergleichsgruppe mit gleichem Experimentaldesign ergänzt (Abschnitt 5.2.4).

realisieren konnten⁴¹. Somit liegen für die Studie 1 Daten zu drei Messzeitpunkten aus jeweils 17 Schulen der Bundesländer Berlin und Brandenburg im Fach Mathematik in einer Vollerhebung vor. Auf Seiten der Schülerinnen und Schüler wurde im Februar 2007 die gesamte neunte Jahrgangsstufe der jeweiligen Schule getestet und befragt. Im folgenden Schuljahr besuchten diese Schülerinnen und Schüler die zehnte Jahrgangsstufe, in der sie zu Beginn (Oktober 2007) und zum Ende des Schuljahres (April 2008) den Test und den Fragebogen erneut vorgelegt bekamen.

Da die Schülerinnen und Schüler für eine weitere Datenerhebung durch das Verlassen der Schulen nicht mehr zur Verfügung standen, wurden für den Messzeitpunkt im Oktober 2008 die zu diesem Zeitpunkt in der neunten Jahrgangsstufe befindlichen Jugendlichen getestet und befragt. Mit diesen Schülerinnen und Schülern wurde im April 2009 auch die Abschlusserhebung durchgeführt.

Hinsichtlich der Lehrkräfte galt es, eine möglichst hohe Teilnahmerate zu erzielen. Daher richtete sich der Lehrerfragebogen nicht nur an die Lehrkräfte der getesteten Klassen, sondern an die gesamte Fachkonferenz Mathematik.

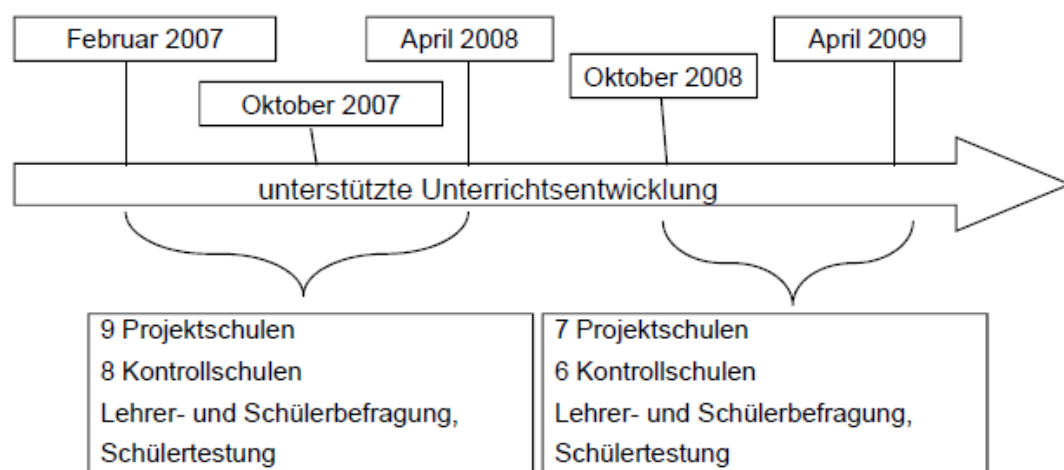


Abbildung 12. Projektdauer und teilnehmende Schulen im Projekt "Erprobung eines Konzepts zur Implementation der Bildungsstandards im Fach Mathematik in Berliner und Brandenburger Schulen"; Quelle: eigene Darstellung.

5.1.5 Stichprobe der Studie 1

Die Stichprobe der Studie 1 wurde durch eine Anfrage bei den für die Bundesländer Berlin und Brandenburg zuständigen Kultusministerien gewonnen. Aus Gründen der Test-

⁴¹ Genaue Gründe für das Ausscheiden wurden von den Schulen nicht benannt.

und Befragungsökonomie sollten für das Projekt möglichst Schulen gewonnen werden, die

1. mindestens dreizügig in der neunten Jahrgangsstufe organisiert sind,
2. über eine gute Anbindung an die öffentlichen Verkehrsmittel verfügen,
3. keine Schulentwicklungserfahrung in Form der Teilnahme an Modellprojekten aufweisen und
4. die Bereitschaft zu einer mindestens einjährigen Teilnahme am Projekt bekunden.

Die von den Ministerien aufgeführten Schulen wurden nach dem Zufallsprinzip vom IQB und LISUM kontaktiert. Die endgültige Bereitschaft zur Teilnahme am Projekt erfolgte nach einer einführenden Projektdarstellung und einem Fachkonferenzbeschluss in den Schulen. Als Quasi-Experiment konzipiert, wurden in gleichem Umfang Schulen akquiriert, die keine Unterstützung in der Entwicklungsarbeit erhielten und folglich auch nicht an inhaltlichen Modulen innerhalb der Fachkonferenzen arbeiteten. Diese nicht-zufällige Schulauswahl umfasst insgesamt vier Gymnasien und fünf Real-, Haupt oder Gesamtschulen aus Berlin sowie weitere drei Gymnasien und fünf Gesamt- bzw. Oberschulen aus dem Bundesland Brandenburg.

Im Zeitraum der Studie 1 wurde je Schulhalbjahr eine Vollerhebung der Fachgruppe Mathematik Sek. I und der sich im Februar 2007 in der neunten Jahrgangsstufe befindenden Schülerinnen und Schüler durchgeführt. Sowohl den Lehrkräften als auch den Schülerinnen und Schülern wurden zu drei Messzeitpunkten die gleichen Fragebögen vorgelegt. Während die Lehrkräfte zu ihrer Arbeit in der Fachkonferenz, ihrer Unterrichtsgestaltung und ihren Einstellungen befragt wurden, sollten die Schülerinnen und Schüler Auskunft zu motivationalen, ihrem Unterrichtsempfinden und Informationen zu ihrem Hintergrund erteilen. Zusätzlich wurden die Schülerinnen und Schüler in einem zweistündigen standard-basierten Mathematiktest getestet. Zu allen drei Messzeitpunkten nahmen insgesamt 67 Klassen in 17 Schulen teil. Tabelle 4 gibt Auskunft über die Klassen- und Schulzuteilung im Experimentaldesign.

Tabelle 4. Übersicht der Schulen und Klassen der Studie 1

Bundesland	Schulform	Anzahl Schulen	Anzahl Klassen		Klassen gesamt
			Projektgruppe	Vergleichsgruppe	
Berlin	Gymnasium	4	13	4	17
	Nicht-Gymnasium	5	7	12	19
Brandenburg	Gymnasium	3	5	6	11
	Nicht-Gymnasium	5	12	7	19
	Gesamt	17	37	29	66

Die längsschnittliche Anlage der Studie führte zu Schwankungen in der Teilnehmeranzahl der einzelnen Messzeitpunkte sowohl auf Lehrer- als auch auf Schülerseite.

An der ersten Erhebung im Februar/März 2007 nahmen 113 *Mathematik-Lehrkräfte* teil (70 weiblich, 40 männlich, 3 ohne Angabe). Davon konnten 65 Personen als Lehrkräfte der Projektschulen identifiziert werden und 48 als Lehrkräfte der Vergleichsschulen. 83.2 Prozent der Lehrkräfte, die hinsichtlich des Alters eine Angabe machten, befanden sich in der Altersspanne zwischen 40 und 59 Jahren⁴². Die zweite Erhebung erfolgte im November 2007 im Schuljahr 2007/2008. Hier waren 82 Mathematiklehrkräfte bereit, den Fragebogen auszufüllen (51 weiblich, 26 männlich, 5 ohne Angabe, 85.4 Prozent zwischen 40 und 59 Jahren). Zum dritten Messzeitpunkt im April 2008 bearbeiteten 73 Mathematiklehrkräfte den Fragebogen (49 weiblich, 18 männlich, 6 ohne Angabe). In der am häufigsten besetzten Altersspanne von 40 bis 59 Jahren konnten hier 85.0 Prozent verortet werden. Abbildung 13 zeigt die Stichprobenverteilung zu allen drei Messzeitpunkten aus der Perspektive der Unterteilung in Projekt- und Vergleichsgruppen.

⁴² 2,7 Prozent der Lehrkräfte machten keine Angaben über ihr Alter.

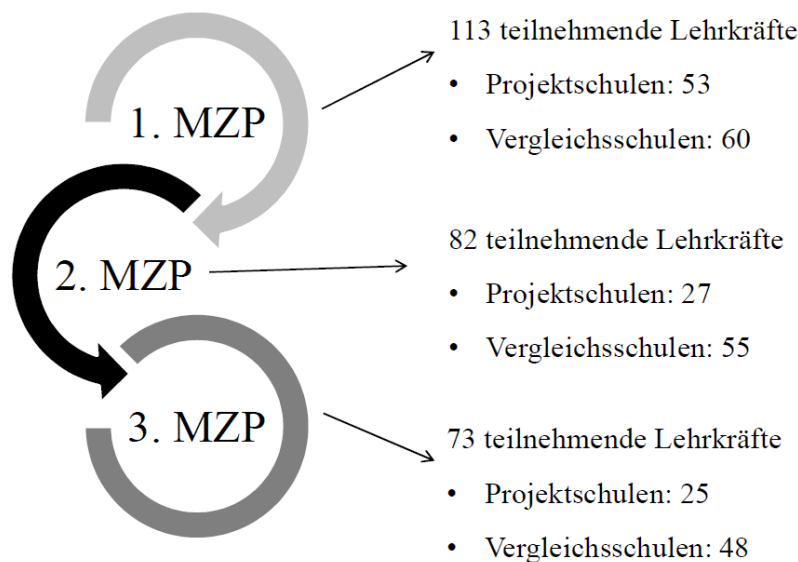


Abbildung 13. Stichprobe der Lehrkräfte der Studie 1.

Der Gesamtdatensatz bestand aus 152 Mathematiklehrkräften, von denen für 44 Personen ein vollständiger Längsschnittdatensatz über drei Messzeitpunkte vorliegt. Diese 44 Personen teilen sich in einem Verhältnis von 16 zu 28 auf die Projekt- und Vergleichsschulen auf.

Weiterhin bearbeiteten 1349 Neuntklässler (709 weiblich, 640 männlich) den Mathematiktest der ersten Erhebung. Die Schülerinnen und Schüler wurden ebenso im Oktober 2007 und April 2008 getestet, als sie sich bereits in der 10. Klassenstufe befanden. Zum zweiten Messzeitpunkt wurden 1106 Schülerinnen und Schüler untersucht, die sich im Verhältnis männlich zu weiblich als 490 zu 608 darstellten (8 ohne Angabe). Zum dritten Messzeitpunkt füllten 1322 Zehntklässler (683 weiblich, 572 männlich, 67 ohne Angabe) unsere Materialien aus.

Gültige Leistungswerte zu allen drei Messzeitpunkten liegen von insgesamt 762 Schülerinnen und Schüler vor. Zum ersten Messzeitpunkt war der Großteil dieser Schülerinnen und Schüler zwischen 15 und 16 Jahren alt (48,5 Prozent mit dem Geburtsjahr 1991 und 40 Prozent mit dem Geburtsjahr 1992).

5.2 Beschreibung der quantitativ-standardisierten Begleitforschung in Hessen

Die quantitativ-standardisierte Begleitforschung zur Implementation der Bildungsstandards wurde am IQB mit dem Schuljahr 2008/2009 erweitert, um auf die Entwicklungen in einzelnen Bundesländern zu reagieren. Während die Schulen in Berlin/Brandenburg weiterhin im Projekt „Erprobung eines Konzeptes zur Implementation der Bildungsstandards im Fach Mathematik in Berliner und Brandenburger Schulen“ verankert waren,

wurden im Bundesland Hessen Schulen für eine zweite Vergleichsgruppe akquiriert. Das in Hessen verfolgte Konzept zur Implementation der Bildungsstandards stützt sich auf die Teilnahme von Fachgruppen bzw. -konferenzen an dem Programm zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ (SINUS). Diesem modularartig aufgebauten Unterrichtsentwicklungsprogramm wird in der Schulentwicklungsarbeit eine Vorreiterrolle zugeschrieben, die ebenso für das Konzept des IQB (Abschnitt 5.1) Pate stand. In der Konzeption des Programms wird ausdrücklich darauf hingewiesen, dass es bei SINUS nicht um „...die Erprobung und anschließende Implementation neuer Unterrichtsansätze geht, sondern um eine Weiterentwicklung des Unterrichts durch die Lehrkräfte an der Basis...“ (Hertrampf, 2004). In Netzwerken organisierte Fachkonferenzen fokussieren ihre Arbeit auf den Umgang mit einer neuen Aufgabenkultur und die Entwicklung eines qualitativ wertvollen Mathematikunterrichts. Konzeptionell folgten dem SINUS-Programm verschiedene Folgeprogramme, deren inhaltliche Schwerpunkte mit den Entwicklungen im Bildungssystem korrespondieren. So wurde mit SINUS-Transfer ein Programm aufgelegt, das in zwei Programmwellen (2003 - 2005 und 2005 - 2007) auf die nachhaltige und flächendeckende Nutzung der SINUS-Ergebnisse sowie deren Weiterentwicklung zielt. Mit dem Folgeprogramm SINUS-Transfer Grundschule (ab 2004) wurden sowohl das Konzept als auch die inhaltlichen Schwerpunkte auf die Grundschule übertragen. In einer dritten Welle der Programmfortführung implementieren die Bundesländer Inhalte aus SINUS-Transfer in neue Projekte, wobei die Organisation dezentral in der Verantwortung der einzelnen Länder liegt. Das Bundesland Hessen bietet seit dem Jahr 2007 das Projekt „Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften“ (KOU) an, das die Kompetenzorientierung im Unterricht, beim Diagnostizieren und Fördern sowie die Implementierung der Bildungsstandards an den Schulen zum Inhalt hat (Kubina, 2009). Aufbauend auf den Erkenntnissen aus dem SINUS-Programm lassen sich in dem aktuellen Programm „Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften“ (KOU) viele konzeptionelle und inhaltliche Parallelen finden, weshalb im Folgenden relevante Aspekte beider Programme vorgestellt werden.

5.2.1 Projektbeschreibung

Das Programm zur Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts wurde als Folge der TIMS-Studie bereits 1998 konzipiert und in 180 Schulländerübergreifend in der Zeit von 1998/1999 bis März 2003 durchgeführt. Die Idee, den

mathematisch-naturwissenschaftlichen Unterricht weiterzuentwickeln, die Professionalisierung der Lehrkräfte voranzutreiben und die Kompetenz- und Interessenentwicklung der Schülerinnen und Schüler zu fördern, entstand durch die in der internationalen Vergleichsstudie sichtbar gewordenen Defizite der deutschen Schüler und wurde durch die ersten PISA-Ergebnisse weiter bestätigt.

Prenzel (2005) führt für den SINUS-Ansatz die Ziele a) der Professionalisierung der Lehrkräfte, b) der Gestaltung des Unterrichts und c) der Förderung von Lernprozessen der Schülerinnen und Schüler an, die sich in einem vertieften, verständnisorientierten Lernen der Schülerinnen und Schüler niederschlagen sollen. Basisorientierung, Nachhaltigkeit und Breitenwirkung werden als Modellversuchsphilosophie ausgelegt (Hertramp, 2004). Mit dem BLK-Programm soll nicht nur die Weiterentwicklung des Unterrichts angestrebt werden, sondern auch eine dauerhafte Etablierung von Qualitätsentwicklungsverfahren in den Fachgruppen der teilnehmenden Schulen stattfinden. Die Entwicklungsrichtung der Fachgruppen wurde mit einem modularen Konzept bestimmt, d.h. die Fachgruppen bearbeiteten elf aufeinander abgestimmte Bausteine, die in ihrer Zusammenstellung auf dem Erkenntnisstand der allgemeinen und fachbezogenen Lehr-Lernforschung beruhen. „Zusammen genommen zielen die elf Bausteine auf die Entwicklung einer deutlich veränderten Kultur des mathematisch-naturwissenschaftlichen Unterrichts, die bedeutungsvolles Lernen, fachliches Verständnis und motivationale Regulierung stärker fördert und fordert.“ (Prenzel et al., 2002, S. 7). Die Autoren des Gutachtens zur Vorbereitung des SINUS-Programms verfolgen damit eine systematische Zusammenarbeit der Lehrkräfte im Fach, der Fächergruppen und bei Bedarf auch darüber hinaus. „Diese Zusammenarbeit erhält ihren größeren Rahmen im regionalen Erfahrungsaustausch zwischen Schulen und in der überregionalen Unterstützung und Betreuung.“ (BLK, 1997, Kap. 9.1). Als Organisationsform etablierte sich im SINUS-Modellversuchsprogramm die Kooperation von jeweils sechs Schulen in lokalen Netzen zu einem Set, wobei eine Schule als Pilotschule fungierte und fünf Schulen die Netzwerkschulen bildeten. Die Pilotschulen fungieren dabei als Verteiler des Programms auf regionaler und landesweiter Ebene und kommunizieren damit die Ideen und Erfahrungen der im Set befindlichen Schulen. Die inhaltliche Arbeit in den Schulen basierte auf Bearbeitung von frei wählbaren Modulen aus einem Angebot von insgesamt elf Modulen.

Im Folgeprogramm KOU wurden Schulsets mit sieben Schulen gebildet, wobei jede Schule von zwei Fortbildnern/innen betreut wird. Zusätzlich werden von den Setlei-

tern/innen prozessorientierte Fortbildungsangebote in die Schulen gegeben, wodurch die fachlichen Fortbildungen ergänzt und ein schulinterner sowie regionaler Erfahrungsaustausch unterstützt wird. Als direkte/r Ansprechpartner/in steht den Lehrkräften ein Schulkoordinator/in aus dem eigenen Kollegium zur Verfügung. Diese Position innerhalb der Organisation des KOU-Projektes umfasst die unmittelbare Organisation, Koordination und Dokumentation der Entwicklungsarbeit an den Schulen. Durch die inhaltlich geänderte Schwerpunktsetzung der Kompetenzorientierung wurden die KMK-Bildungsstandards zu den Zielen von Unterricht. Die Durchführung des Programms lehnt sich dabei stark an die für das SINUS-Programm geltenden Leitsätze an (Kubina, 2009, BLK, 1997):

- regelmäßige Teilnahme an den fachlichen und prozessbegleitenden Fortbildungen,
- kontinuierliche Arbeit in den Teams an der Weiterentwicklung des Unterrichts mit dem Ziel der Kompetenzförderung der Schülerinnen und Schüler (z.B. Entwicklung kompetenzorientierter Aufgaben, Unterstützung des selbstverantwortlichen Lernens durch Lernstandsanalyse und Förderung),
- Orientierung an mathematischer/naturwissenschaftlicher Grundbildung,
- kollegiale Kooperation innerhalb des Fachkollegiums,
- kontinuierlicher Erfahrungsaustausch der am Set teilnehmenden Schulen (Kubina, 2009, S. 38).

5.2.2 Modulbeschreibungen

Die Module stellen den Schwerpunkt des Qualitätsentwicklungskonzeptes dar und sollen daher im Folgenden inhaltlich beschrieben werden:

1. Weiterentwicklung der Aufgabenkultur
2. Naturwissenschaftliches Arbeiten
3. Aus Fehlern lernen
4. Sicherung von Basiswissen – verständnisvolles Lernen auf unterschiedlichen Niveaus
5. Zuwachs von Kompetenz erfahrbar machen: kumulatives Lernen
6. Fächergrenzen erfahrbar machen – fachübergreifendes und fächerverbindendes Arbeiten
7. Förderung von Mädchen und Jungen
8. Entwicklung von Aufgaben für die Kooperation von Schülern

9. Verantwortung für das eigene Lernen stärken
10. Prüfen: Erfassen und Rückmelden von Kompetenzzuwachs
11. Qualitätssicherung innerhalb der Schule und Entwicklung schulübergreifender Standards.

Die Module wurden im Verlauf der Programmarbeit inhaltlich durch Erläuterungen oder Beispielbeschreibungen konkretisiert. Im Disseminationskonzept des Programms wird jedoch festgehalten, dass die entstandenen Materialsammlungen und Handreichungen nicht per se übernommen werden können, vielmehr sind zukünftig im Programm arbeitende Lehrkräfte angehalten, eigene Überlegungen zu entwickeln, um die individuellen Startpunkte für eine problemorientierte Unterrichtsentwicklung zu identifizieren (IPN, 2002). In den Folgeprojekten dienten die entstandenen Materialien jedoch als Grundlage der Weiterentwicklung. Alle im Anschluss an das Modellversuchsprogramm SINUS konzipierten Projekte übernahmen die Idee der Modulkonzeption und die Kooperation in Netzwerken. Inhaltlich verwendeten die Nachfolgeprojekte SINUS-Transfer die Module aus dem Modellversuchsprogramm und führten immer mehr Schulen an die qualitätsentwickelnden Maßnahmen heran. Mit dem KOU-Projekt reagierte das Bundesland Hessen auf die nationalen Bildungsstandards. Die erste Welle dieser Fortbildungsreihe arbeitete nach dem Konzept des SINUS-Modellversuchsprogramms mit Schulen in der Zeit von 2007 bis 2009 mit der bestehenden Leitlinie der kooperativen Qualitätsentwicklung von Unterricht durch die Lehrkräfte. Die teilnehmenden Schulen wurden hier jedoch zur Bearbeitung dreier Module verpflichtet und für die verbleibenden fünf Module aufgefordert, eine Auswahl zu treffen. Auch in dieser Fortbildungsreihe wurden die verfügbaren Materialien aus den vergangenen SINUS-Wellen zur Weiterentwicklung des Unterrichts mit dem Ziel der Kompetenzförderung von Schülerinnen und Schülern genutzt. Abbildung 14 unterscheidet die insgesamt acht Module des Faches Mathematik in obligatorische und fakultative Angebote.

MATHEMATIK	NATURWISSENSCHAFTEN
PFLICHTMODULE	
1. MODUL: PERSPEKTIVWECHSEL Vom Umgang mit Fehlern zum Umgang mit Bildungsstandards	1. MODUL: ANNÄHRUNG AN DIE KOMPETENZORIENTIERUNG IN DEN NATURWISSENSCHAFTEN Grundsätze verständnisvollen Lehrens und Lernens (Vom Defizit zum Stärkmodell, Umgang mit Fehlern), Bildungsstandards – Kompetenzen – Kompetenzstufen – Basiskonzepte NaWi
2. MODUL: AUFGABEN Mit kompetenzorientierten Aufgaben zu den Bildungsstandards	2. MODUL: KOMPETENZORIENTIERTER UNTERRICHT Mit kompetenzorientierten Unterrichtssituationen naturwissenschaftliches Denken und Arbeiten anregen
3. MODUL: DIAGNOSE UND FÖRDERUNG Pädagogische Diagnostik – individuelle Kompetenzförderung und Bildungsstandards	3. MODUL: DIAGNOSE UND FÖRDERUNG Pädagogische Diagnostik – individuelle Kompetenzförderung und Bildungsstandards, Kompetenzermittlung und Aufgabenentwicklung
WAHLMODULE	
MODUL: KOMPETENZORIENTIERTE UNTERRICHTSSEQUENZEN PLANEN UND DURCHFÜHREN Förderkreislauf, Zielorientierung in einer kompetenzorientierten Unterrichtssequenz	MODUL: ERKENNTNISGEWINNUNG Naturwissenschaftliche Denk- und Arbeitsweisen
MODUL: MATHEMATIK UND SPRACHE Aspekte: Methoden zur Erhöhung der individuellen Sprachanteile im MU; Verbesserung des Text- Leseverständnisses; von der Alltagssprache zur Fachsprache	MODUL: KOMMUNIKATION Von der Alltagssprache zur Fachsprache – die verschiedenen Abstraktionsebenen von Sprache
MODUL: BEURTEILEN UND BEWERTEN Beurteilung und Bewertung in offenen Unterrichtssituationen	MODUL: KOMMUNIZIEREN UND BEWERTEN NATURWISSENSCHAFTLICHER INFORMATIONEN Kompetenzbereich Kommunikation und Bewertung
MODUL: MIT LERNUMGEBUNGEN ARBEITEN Entwickeln von Grundvorstellungen	MODUL: LEISTUNGSRÜCKMELDUNG ... in offenen Lernsituationen, bei selbst bestimmten Arbeitsformen, in kompetenzorientierten Lernumgebungen
MODUL: VERTIEFUNG DIAGNOSE UND FÖRDERN Diagnose- und Fördermaßnahmen, Lernvereinbarungen	MODUL: NATURWISSENSCHAFTEN IN 5/6 Kompetenzorientiert unterrichten im Lernbereich Naturwissenschaften im Jg. 5/6
	MODUL: MIT METHODEN KOMPETENZEN FÖRDERN Methodenwerkzeuge als Instrumente zur Kompetenzförderung

Abbildung 14. Übersicht über die Module im Projekt "Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften"; Quelle: Kubina, 2009.

5.2.3 Durchführung der Begleitforschung am IQB

Die Testungen und Befragungen in den Bundesländern Berlin und Brandenburg erfolgten in der Frequenz eines Schulhalbjahres, wodurch bei den Lehrkräften ein routinierter Umgang mit externen Testleitern erfolgte. Für die Schulen dieser Stichprobe fand das Konzept des IQB zur Implementation der Bildungsstandards im Fach Mathematik weiterhin Anwendung. In diesem für die Schulen zweiten Projektjahr fanden neben der fortzuführenden Arbeit an den Modulen zusätzlich zwei prozessbegleitende Fortbildungen statt (Tabelle 3).

Zur Durchführung der Testungen im Bundesland Hessen wurden in Kooperation mit der Universität Kassel und der Universität Göttingen Testleiter akquiriert und vor Ort eine umfassende Testleiterschulung veranstaltet. Die Testleiter wurden dabei mit den Besonderheiten der Studie vertraut gemacht und auf die Beachtung der Gütekriterien einer standardorientierten Testung hingewiesen. Die Schulen konnten durch eine Anfrage beim Landesinstitut für Lehrerbildung Hessen bestimmt werden und wurden vom IQB telefo-

nisch angefragt. Nach einem Entschluss der Fachkonferenz zur Teilnahme an der Begleitforschung erfolgte eine zeitnahe Testung und Befragung der Schulen. Die Schulen erhielten nach jeder Testung eine schriftliche Rückmeldung zu den Schülerleistungen, die eine Einordnung der Schülerkompetenzen in einen nationalen Leistungsvergleich ermöglichte⁴³.

5.2.4 Stichprobe der Studie 2

Die Stichprobe der Studie 2 umfasst auf Schulebene 13 Fachgruppen aus dem Raum Berlin/Brandenburg und weitere 17 Fachkonferenzen aus dem Bundesland Hessen. Während sich die Stichprobe aus Berlin/Brandenburg auf Schulebene verkleinerte⁴⁴, konnten in Hessen weitere Schulen für die Begleitforschung gewonnen werden. Die Schulakquise erfolgte nach den gleichen Prinzipien wie in Berlin/Brandenburg.

⁴³ Die rückgemeldeten mathematischen Kompetenzen wurden im Vergleich zu den Ergebnissen der nationalen Normierungsstudie dargestellt.

⁴⁴ Ein Gymnasium aus dem Bundesland Berlin entschied sich gegen eine Fortführung des Projektes. Eine Gesamtschule aus dem Bundesland Brandenburg konnte aufgrund zu geringer Schülerzahlen keine Neuntklässler mehr unterrichten, der Jahrgang wurde in einer anderen Schule eingegliedert. Die Fachgruppe sah dadurch keine Verpflichtung mehr darin, am Projekt noch weiter teilzunehmen. Weiterhin konnten in zwei Vergleichsschulen keine Daten mehr gesammelt werden; die Schulen argumentierten mit einem unausgewogenen Aufwand-Nutzen-Verhältnis der Schülertestung und –befragung.

Tabelle 5. *Schulentwicklungserfahrung und Zuordnung im Experimentaldesign der hessischen Schulen*

	Teilnahme an SINUS 2003- 2005	Teilnahme an SINUS 2005- 2007	Teilnahme an KOU (Mathe) ab 2007	Gruppenzuordnung
Schule 1	-	X (Nawi)	-	Vergleichsgruppe
Schule 2	-	X (Mathe)	X	Projektgruppe
Schule 3	-	X (Mathe)	X	Projektgruppe
Schule 4	-	X (Mathe)	X	Projektgruppe
Schule 5	-	X (Mathe)	X	Projektgruppe
Schule 6	X	-	X	Projektgruppe
Schule 7	X	-	X	Projektgruppe
Schule 8	-	-	-	Vergleichsgruppe
Schule 9	-	-	X	Projektgruppe
Schule 10	X	X (Nawi)	-	Vergleichsgruppe
Schule 11	-	X (Mathe)	-	Vergleichsgruppe
Schule 12	-	X (Mathe)	X	Projektgruppe
Schule 13	-	X (Nawi)	X	Projektgruppe
Schule 14	-	-	X	Projektgruppe
Schule 15	-	-	-	Vergleichsgruppe
Schule 16	-	X (Mathe)	-	Vergleichsgruppe
Schule 17	-	X (Mathe)	-	Vergleichsgruppe

Das Kriterium für die Zuordnung von Projekt- und Vergleichsgruppe wurde hier durch die Teilnahme an SINUS-Programmen seit dem 01.02.2008 gekennzeichnet.

Die vom damaligen Amt für Lehrerbildung Hessen (AfL)⁴⁵ benannten hessischen Schulen mit Schulentwicklungserfahrung weisen in einigen Fällen zusätzlich zur Teilnahme im aktuellen Projekt KOU eine vorherige Teilnahme im SINUS-Transfer-Projekt auf. Die Schulentwicklungserfahrung der hessischen Stichprobe ist in Anhang A ausführlich dargestellt. Aus Tabelle 5 geht hervor, dass die Zuordnung zu Projekt- oder Vergleichsgruppe aus dem Kriterium der Teilnahme an der Fortbildungsreihe KOU erfolgte. Damit werden die Schulen, die vorher an SINUS-Transfer teilnahmen, ohne das KOU-Projekt fortzuführen, der Vergleichsgruppe zugewiesen, obwohl bereits ein gewisser Grad an Schulentwicklungserfahrung vorliegt. Aufgrund der inhaltlich unterschiedlichen Schwerpunktsetzungen in den Programmen SINUS-Transfer und KOU wiegen die Informationen und

⁴⁵ Das hessische Amt für Lehrerbildung ist 2013 in das „Landesschulamt und Lehrkräfteakademie“ übergegangen.

Kenntnisse zur Kompetenzorientierung aus dem KOU-Projekt bei der Gruppenzuordnung in der Begleitforschung des IQB schwerer, weshalb die vorherige Teilnahme der Schulen 11 und 16 an SINUS-Transfer unberücksichtigt bleibt.

Auf Schul- bzw. Klassenebene stellt sich die Stichprobe der zweiten Studie im quasi-experimentellen Design wie in Tabelle 6 dar.

Tabelle 6. Übersicht der Schulen und Klassen in Studie 2

Bundesland	Schulform	Anzahl Schulen	Anzahl Klassen		Klassen gesamt
			Projektgruppe	Vergleichsgruppe	
Berlin	Gymnasium	2	7	0	7
	Nicht-Gymnasium	5	7	13	20
Brandenburg	Gymnasium	3	3	4	7
	Nicht-Gymnasium	3	5	4	9
Hessen	Gymnasium	8	23	15	38
	Nicht-Gymnasium	9	16	16	32
	Gesamt	30	61	52	113

Die Studie 2 besteht aus zwei Messzeitpunkten mit jeweils 113 Klassen. Die Lehrer- und Schülerbefragung erfolgte mit den Instrumenten der Studie 1. Die befragte und getestete Schülerkohorte bestand hier aus den Schülerinnen und Schülern der neunten und zehnten Jahrgangsstufe, wobei in den Bundesländern Berlin und Brandenburg im Oktober 2008 die Neuntklässler getestet und befragt wurden, während sich der Testzeitraum in Hessen durch Verzögerungen in der Schulakquise von Mai 2008 bis Oktober 2008 erstreckte. Die von Mai bis Juli getesteten Schülerinnen und Schüler in Hessen (Schuljahr 2007/2008) befanden sich in der neunten Jahrgangsstufe (19 Klassen), die verbleibenden Schülerinnen und Schüler in 51 Klassen aus dem Bundesland Hessen werden als Zehntklässler charakterisiert (Schuljahr 2008/2009).

Aufgrund der inhaltlichen und konzeptionellen Ähnlichkeit der beiden Interventionen des IQB-Implementationskonzeptes und des KOU-Programms werden die Personen beider Interventionsformen im Weiteren als eine gemeinsame Projektgruppe betrachtet. Die verbleibenden Schulen werden gemeinschaftlich als Kontrollgruppe bezeichnet.

Zum ersten Messzeitpunkt liegen Daten von 171 Lehrkräften vor. Die Lehrerinnen bilden darin eine Gruppe von 89 Personen (52 Prozent, 77 männlich, 5 ohne Angabe). Der Anteil der Lehrkräfte im Alter von 40 bis 59 Jahren liegt bei 69 Prozent. Von diesen Lehrkräften gaben 93,4 Prozent an, das Fach Mathematik studiert zu haben.

Zum zweiten Messzeitpunkt nahmen 135 Lehrkräfte an der Befragung teil (113 weiblich, 17 männlich, 5 ohne Angabe). Mit einem Anteil von 60 Prozent bilden auch hier die Personen zwischen 40 und 59 Jahren die Mehrheit der Teilnehmer.

In der Schülerstichprobe können zum ersten Messzeitpunkt gültige Leistungswerte von 2.458 Schülerinnen und Schülern abgerufen werden. Das Geschlechterverhältnis männlich zu weiblich gestaltet sich in den Teilnahmegrößen von 1.139 zu 1.264 (55 ohne Angabe).

Die zweite Datenerhebung verzeichnete Ergebnisse von 2.151 Schülerinnen und Schülern. Davon waren 1.133 weiblichen Geschlechts und 992 männlichen Geschlechts (26 ohne Angabe). Gültige Leistungswerte für beide Testungen liegen von 1.970 Schülerinnen und Schülern vor (1.061 bzw. 53.9 Prozent weiblich, 906 bzw. 46.1 Prozent männlich, 3 ohne Angabe).

5.3 Die eingesetzten Instrumente

Die hier vorgestellten Instrumente zur Messung relevanter Aspekte der Implementation der Bildungsstandards wurden in beiden Studien durchgehend verwendet, ohne Veränderungen oder Kürzungen der Skalen erfahren zu haben. Die ausführlichen Itemformulierungen sind in Anhang A einsehbar. Die Fülle der im Gesamtfragebogen erfassten Konstrukte und Aspekte der Bildungsstandards wurde von den Lehrkräften häufig kritisiert und ist als eine Ursache für die sinkenden Teilnahmeraten im Längsschnitt nicht auszuschließen (Abschnitt 6.3).

Die Güte der vorgestellten Instrumente wird jeweils durch das Reliabilitätsmaß Cronbach's Alpha und mit Blick auf die Mehrebenenstruktur der Daten durch die Intraklasssenkorrelationsmaße ICC (1) und ICC(2) (Bliese, 2000; Lüdtke et al, 2006; Raudenbush & Bryk, 2002) bestimmt. Mit den Indizes ICC (1) und der ICC(2) werden Maße bereitgestellt, die auf einer einfaktoriellen Varianzanalyse basieren und die Wahrnehmung der Schülerinnen und Schüler hinsichtlich eines bestimmten Merkmals quantifizieren. Beide Reliabilitätsmaße können Werte zwischen 0 und 1 annehmen, wobei der ICC (1) den prozentuellen Anteil der Gesamtvarianz angibt, der auf Unterschiede zwischen den Klassen zurückzuführen ist und der ICC (2) ein Maß für die Genauigkeit des über alle Schüler einer Klasse gemittelten Urteils darstellt (Lüdtke et al., 2006, S. 87). Die Dokumentation der Kennwerte ICC (1) und ICC (2) erfolgt für alle eingesetzten Fragebögen. Die Kennwerte werden im Querschnitt für den jeweiligen Messzeitpunkt betrachtet. Als Gruppen-

variable wird bei den Lehrkräften die jeweilige Fachkonferenz der Schule betrachtet, bei den Schülerinnen und Schülern die jeweilige Klasse.

5.3.1 Einstellungen gegenüber den Bildungsstandards

Der von Freudenthaler und Specht (2004) entwickelte Fragebogen zu den allgemeinen Einstellungen gegenüber den Bildungsstandards umfasst insgesamt 23 Items und wurde in der ersten Pilotphase zur Implementation nationaler Bildungsstandards in Österreich eingesetzt. Der Intention folgend, den Zustimmungsgrad von Lehrkräften zu elf positiven und zwölf negativen Aspekten der Einführung von Bildungsstandards zu erfragen, konzipierten die Forscher eine 5-stufige Ratingskala von 1 stimmt genau bis 5 stimmt gar nicht. Der österreichischen Intention folgend, wurden aus dem Gesamtfragebogen zwei Skalen extrahiert, die einerseits die positiven Aspekte bei der Einführung nationaler Bildungsstandards abbilden und andererseits die negativen Aspekte. Dadurch konnten alle Items in ihrer ursprünglichen Kodierung verwendet werden, womit höhere Antwortwerte mit einer stärkeren Ausprägung des Konstrukts assoziieren⁴⁶. Tabelle 7 führt beispielhaft zwei Items aus dem Fragebogen auf.

Tabelle 7. *Beispielitems für die Einstellungsskalen zur Einführung der nationalen Bildungsstandards*

Skala	Beispielitem
Positive Einstellungen	Bildungsstandards schärfen den Blick der Lehrkraft für Stärken und Schwächen ihrer Schüler/innen.
Negative Einstellungen	Bildungsstandards bringen die Gefahr einer inhaltlichen Verarmung des Unterrichts mit sich.

In Tabelle 8 werden zusammengefasst die Reliabilitätskennwerte der Skalen aufgeführt. Beide Skalen weisen eine hohe innere Konsistenz mit Werten von $\alpha = .82$ bis $\alpha = .91$ auf. Die Werte des ICC (1) weisen in einem Umfang von $ICC(1) = .23$ bis $ICC(1) = .48$ darauf hin, dass es nur moderate Übereinstimmungen in den Urteilen der Lehrkräfte einer Fachkonferenz gibt.

Mit einer in der Literatur berichteten akzeptablen Grenze von .70 für Werte des ICC(2) als Reliabilitätsmaß der Klassenmittelwerte weisen die hier berechneten ICC(2)-Koeffizienten mehrheitlich Werte größer .70 auf. Der geringste Wert von .60 ergibt sich aus den Daten der Studie 1 zum zweiten Messzeitpunkt und weist darauf hin, dass die Urteile der Lehrkräfte hier stärker variieren und damit eine Verunsicherung in der Einstellungsfrage gegenüber den Bildungsstandards stattfand, was auf eine Auseinandersetzung mit der Thematik hinweist.

⁴⁶ Die ausführlichen Fragebögen sind in Anhang A nachzulesen.

Tabelle 8. *Reliabilitätsmaße der Skalen zu den Einstellungen gegenüber den Bildungsstandards (Lehrkräfte)*

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
positive Einstellungen	.90/ .26/ .70	.90/ .24/ .60	.85/ .37/ .73	.88/ .31/ .71	.91/ .48/ .81
negative Einstellungen	.86/ .23/ .67	.84/ .26/ .63	.82/ .29/ .66	.88/ .40/ .78	.90/ .38/ .74

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

5.3.2 Kooperation und Zusammenarbeit in der Fachkonferenz

Zur Erfassung des latenten Konstrukts „Kooperation“ wurden zwei Skalen eingesetzt, die bereits in der PISA 2006-Erhebung Verwendung fanden (Frey et al., 2009). Die Skala „Zusammenarbeit in der Fachkonferenz“ beschreibt mit einer Antwortskala von 1 (nie) bis 6 (jeden Tag), wie häufig bestimmte Themen und Inhalte in den Fachkonferenzen der Mathematiklehrkräfte besprochen werden. Die Skala umfasst insgesamt acht Items, die alle positiv formuliert wurden und das Ende der Aussage „In der Fachkonferenz...“ bilden, z.B. „...beraten wir die Möglichkeiten individueller Förderung für Schülerinnen und Schüler.“. Jeweils ein Item aus den Fragebögen ist in Tabelle 9 abgebildet. Der ausführliche Fragebogen ist in Anhang A einsehbar.

Tabelle 9. *Beispielitems zu den Skalen Zusammenarbeit in der Fachkonferenz und Kooperation*

Skala	Beispielitem
Zusammenarbeit in der Fachkonferenz	In der Fachkonferenz ...treffen wir uns, um jahrgangsübergreifende Unterrichtsplanung auszuarbeiten.
Kooperation in der Fachkonferenz	Bei der Zusammenarbeit in der Fachkonferenz Mathematik hatte ich den Eindruck, dass...in der Gruppe Einigkeit über die Zielvorstellung herrscht.

Die Items betonen mehrheitlich den organisatorischen Teil der Fachkonferenzarbeit und sprechen die überwiegenden Themen an, die die Lehrkräfte einer Fachgruppe miteinander teilen, z.B. Lehrbuchauswahl oder Fortbildungsveranstaltungen.

Die Skala „Kooperation“ fragt hingegen vorrangig nach der Qualität der Zusammenarbeit in der Fachkonferenz. Mit der Beendigung der Aussage „Bei der Zusammenarbeit in der Fachkonferenz hatte ich den Eindruck, dass...“ werden mithilfe von neun Items die At-

mosphäre und Arbeitsweise in der Fachkonferenzarbeit erfragt, z.B. „...die Arbeitsteilung gut gelingt.“, „...wir mit einer klaren Zielvorstellung an die Arbeit gehen.“. Sieben der neun Items waren hinsichtlich der Skala von 1 (trifft nicht zu) bis 4 (trifft zu) positiv formuliert und zwei Items negativ formuliert, z.B. „...ich alleine wesentlich effektiver arbeite.“. Die zwei negativ formulierten Items wurden für die weitere Arbeit rekodiert.

Tabelle 10. Reliabilitätsmaße der Skalen „Zusammenarbeit in der Fachkonferenz“ und „Kooperation in der Fachkonferenz“ (Lehrkräfte)

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
Zusammenarbeit in der FK	.69/ .31/ .74	.82/ .29/ .68	.78/ .42/ .78	.81/ .25/ .65	.78/ .31/ .68
Kooperation in der FK	.91/ .44/ .84	.88/ .41/ .78	.90/ .68/ .91	.91/ .33/ .73	.93/ .51/ .83

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

Mit Ausnahme der Skala zur Zusammenarbeit in der Fachkonferenz zum ersten Messzeitpunkt der Studie 1 weisen die α -Werte auf eine hohe innere Konsistenz der Skalen hin. Sowohl die ICC(1) als auch die ICC(2)-Werte zeigen keine auffälligen Ausreißer.

5.3.3 Kompetenzorientierte Schülertätigkeiten

Der Untersuchungsgegenstand „Kompetenzorientierter Unterricht“ wurde mit einem 40 Items umfassenden Fragebogen untersucht, der die Schülertätigkeiten im Mathematikunterricht erfragt. Die Items wurden dabei aus den Bildungsstandards für das Fach Mathematik Sekundarstufe I abgeleitet. So konnten für jede der sechs allgemeinen mathematischen Kompetenzen fünf bis sieben Items formuliert werden, die eine Tätigkeitsbeschreibung beinhalten, wie sie im Kompetenzmodell in der Dimension der allgemeinen mathematischen Kompetenzen ausgewiesen wird. Der Fragebogen wurde sowohl den Lehrkräften als auch den Schülerinnen und Schülern vorgelegt und erlaubt daher eine Gegenüberstellung der unterschiedlichen Perspektiven auf den Mathematikunterricht. Als Antwortskala wurde eine Häufigkeitsskala von 1-nie bis 6-jeden Tag gewählt. Während die Lehrkräfte dazu aufgefordert wurden, die Aussage „Wie oft sollten die Schülerinnen und Schüler in den vergangenen 6 Monaten in Ihrem Unterricht....“ zu ergänzen, beantworteten die Schülerinnen und Schüler die Frage „Wie oft habt ihr in deiner Klasse in den ver-

gangenen 6 Monaten im Mathematikunterricht Folgendes ausgeführt?“). Durch die inhaltliche Überschneidung einiger Kompetenzen (z.B. argumentieren und kommunizieren) ist eine konfirmatorische Faktorenanalyse für die aus der Theorie formulierten Items nicht sinnvoll. Vielmehr wird hier in einem ersten Schritt angestrebt, eine einheitliche Itemanzahl pro Skala zu extrahieren. Dazu wurden die Items zur Beschreibung einer Kompetenzskala einer Trennschärfeuntersuchung unterzogen und auf fünf Items pro allgemeine Kompetenzskala verkürzt.

Somit ergeben sich sechs Skalen entsprechend den sechs allgemeinen mathematischen Kompetenzen mit jeweils fünf Items, für die in Tabelle 11 jeweils ein Beispielitem abgebildet ist.

Tabelle 11. *Beispielitems zum Fragebogen "Kompetenzorientierte Schülertätigkeiten"*

Skala	Beispielitem
<i>Wie oft sollten die Schülerinnen und Schüler in den vergangenen 6 Monaten in Ihrem Unterricht</i>	
mathematisches Argumentieren	...den Gedankengang erklären, der hinter einer Idee steht.
Probleme lösen	...Aufgaben bearbeiten, bei denen nicht sofort der Lösungsweg zu erkennen ist.
mathematisches Modellieren	...mathematische Strukturen in Alltagskontexten erkennen.
mathematische Darstellungen verwenden	...Beziehungen zwischen Tabellen, Graphen oder Diagrammen erkennen.
technisch Arbeiten	...mathematische Sätze, Regeln oder Formeln anwenden.
mathematisches Kommunizieren	...eigene Lösungswege anderen Schülern erläutern.

Items, die aufgrund einer zu geringen Trennschärfe aus der Skala ausgeschlossen wurden, werden in den weiteren Berechnungen nicht berücksichtigt. Die Skalen wurden aus den Daten des ersten Messzeitpunktes gebildet und sind für die weiteren Messzeitpunkte bindend.

Hinsichtlich einer inhaltlichen Vergleichbarkeit der Skalen, werden die gleichen Items für die Skalenbildung sowohl aus den Daten der Lehrkräfte als auch der Daten der Schülerinnen und Schüler verwendet⁴⁷.

⁴⁷ D.h. Items, die aufgrund einer zu geringen Trennschärfe aus der Skalenbildung zum 1. MZP, Stichprobe Lehrkräfte ausgeschlossen wurden, fanden auch zu späteren Messzeitpunkten und auch in den Daten der Schülerinnen und Schüler keine Verwendung.

Tabelle 12. Reliabilitätsmaße zu den Skalen der kompetenzorientierten Tätigkeiten (Lehrkräfte)

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
argumentieren	.81/ .25/ .69	.76/ .25/ .63	.73/ .28/ .65	.80/ .22/ .62	.76/ .27/ .64
Probleme lösen	.76/ .21/ .64	.76/ .22/ .59	.74/ .30/ .68	.76/ .13/ .46	.77/ .25/ .61
modellieren	.87/ .20/ .62	.89/ .27/ .65	.76/ .32/ .69	.87/ .21/ .60	.86/ .25/ .61
Darstellungen verwenden	.85/ .21/ .63	.81/ .17/ .51	.83/ .33/ .71	.88/ .26/ .67	.84/ .28/ .65
technisch arbeiten	.64/ .29/ .72	.55/ .21/ .57	.64/ .48/ .82	.65/ .25/ .66	.64/ .32/ .69
kommunizieren	.73/ .17/ .57	.82/ .10/ .37	.47/ .27/ .64	.76/ .15/ .49	.79/ .28/ .66

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

Die teilweise recht geringe Reliabilität der Einzelurteile der Lehrkräfte (ICC(1)) weist darauf hin, dass sich der Unterricht der Lehrkräfte innerhalb einer Schule stark voneinander unterscheidet. Die Reliabilitätswerte des ICC(2) rangieren in einem hohen Maß zwischen .37 und .82. Die mehrheitlich geringen Werte des ICC(2) zeigen, wie stark der kompetenzorientierte Mathematikunterricht von der unterrichtenden Lehrkraft abhängt. Für die Schülerinnen und Schüler ergeben sich folgende Werte für die Gütekriterien des Instruments (Tabelle 13).

Tabelle 13. *Reliabilitätsmaße zu den Skalen der kompetenzorientierten Tätigkeiten (Schülerinnen und Schüler)*

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
argumentieren	.75/ .19/ .84	.77/ .24/ .86	.77/ .18/ .83	.75/ .23/ .87	.75/ .26/ .88
Probleme lösen	.68/ .15/ .79	.69/ .19/ .83	.71/ .14/ .78	.70/ .18/ .83	.71/ .21/ .85
modellieren	.78/ .14/ .78	.79/ .17/ .81	.77/ .17/ .82	.78/ .16/ .80	.79/ .16/ .79
Darstellungen verwenden	.74/ .19/ .84	.77/ .26/ .87	.78/ .21/ .85	.73/ .17/ .81	.76/ .19/ .82
technisch arbeiten	.56/ .17/ .82	.54/ .19/ .82	.59/ .16/ .81	.55/ .14/ .78	.58/ .17/ .80
kommunizieren	.68/ .18/ .83	.69/ .22/ .85	.71/ .17/ .82	.68/ .17/ .82	.69/ .18/ .81

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2/(\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1))/(1 + (k - 1) * ICC(1))$

Die Skalen zu den kompetenzorientierten Tätigkeiten aus Schülersicht weisen eine recht geringe innere Konsistenz auf. Ebenso die Intraklassenkorrelation ICC(1) erreicht maximale Werte von .26. Demnach ist die Reliabilität des einzelnen Schülerurteils in der Klasse nicht besonders hoch. Die Unterrichtswahrnehmung aus Schülersicht scheint sehr individuell empfunden zu werden. Dagegen übersteigen alle ICC(2)-Werte die Grenze von .70 und weisen damit eine hohe Reliabilität des Klassenmittelwertes hinsichtlich des zu untersuchenden Merkmals auf.

5.3.4 Mathematisches Selbstkonzept der Schülerinnen und Schüler

Auf Seiten der Schülerinnen und Schüler wurde im Fragebogen auch nach ihrem mathematischem Selbstkonzept gefragt, das als relevantes motivationales Konstrukt einen Einfluss auf die Mathematikleistung hat (Köller, 2004). In Anlehnung an die PISA-2003-Erhebung wurden hier fünf Items eingesetzt, die sich auf die Selbstbetrachtung im Hinblick auf die Mathematik beziehen. Diese Skala fand bereits Einsatz im internationalen Schülerfragebogen zur PISA-2003-Erhebung (Ramm et al., 2006). Fünf weitere Items, die im nationalen Schülerfragebogen in PISA eingesetzt wurden, erfragen das Interesse der Schülerinnen und Schüler an Mathematik. Die je fünf Items werden zu den Skalen *Selbstkonzept* und *Interesse* zusammenfasst.

Alle Items sollten auf einer Zustimmungsskala von 1 (stimmt überhaupt nicht) bis 4 (stimmt ganz genau) beantwortet werden. Das erste Item der Skala *Selbstkonzept* wurde aufgrund seiner negativen Formulierung rekodiert. *Tabelle 14* bildet exemplarisch je ein Item aus den Skalen ab. Eine Gesamtübersicht der Items ist in Anhang A einzusehen.

Tabelle 14. *Beispielitems zu den Skalen "mathematisches Selbstkonzept" und "Interesse an Mathematik" im Schülerfragebogen*

Skala	Beispielitem
Mathematisches Selbstkonzept	Ich war schon immer überzeugt, dass Mathematik eines meiner besten Fächer ist.
Mathematisches Interesse	Die Beschäftigung mit Mathematik gehört zu meinen Lieblingstätigkeiten.

Die innere Konsistenz der beiden Skalen weist mit Werten von .86 bis .91 hohe Reliabilitäten auf. Dagegen zeigen die sehr geringen Werte des ICC(1), dass die Reliabilität eines einzelnen Schülers einer Klasse sehr gering ist. Folglich variiert sowohl das mathematische Selbstkonzept als auch das mathematische Interesse der Schülerinnen und Schüler in einer Klasse stark. Die Unterschiede in den jeweiligen Beurteilungen liegen damit vielmehr in den Klassen als zwischen den Klassen. Die geringen Werte weisen auch auf eine heterogene Klassenzusammensetzung hinsichtlich dieser beiden Merkmale hin. *Tabelle 15* weist ferner den ICC(2) aus, dessen Werte den Grenzwert von .70 nur im zweiten Messzeitpunkt der Studie 1 bei der Beurteilung des mathematischen Selbstkonzeptes übertreffen. Die Reliabilität des Klassenmittelwertes weist damit eher geringe Werte auf.

Tabelle 15. *Reliabilitätsmaße zu den Skalen „mathematisches Selbstkonzept“ und „Interesse an Mathematik“ (Schülerinnen und Schüler)*

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
math. Selbstkonzept	.88/ .09/ .67	.90/ .12/ .73	.90/ .08/ .67	.89/ .06/ .58	.91/ .07/ .59
math. Interesse	.86/ .07/ .64	.87/ .10/ .70	.88/ .08/ .67	.87/ .09/ .69	.88/ .09/ .67

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

5.3.5 Die Wahrnehmung des Lehrers durch die Schülerinnen und Schüler

Zusätzlich zu den motivationalen Aspekten sollten die Schülerinnen und Schüler eine Einschätzung ihres/ihrer Mathematiklehrers/in vornehmen. Auch dieses Merkmal kann ebenso wie die motivationalen Konstrukte kontrollierend für die Bestimmung der Mathematikleistung eingesetzt werden. Die Skala besteht aus sechs Items, wurde für die COACTIV-Studie 2003/2004 entwickelt und fragt auf einer Skala von 1 (trifft überhaupt nicht zu) bis 4 (trifft ganz zu), wie die Schülerinnen und Schüler ihre/n Mathematiklehrerin/er erlebt haben. Von den insgesamt sechs Items wurde ein Item negativ formuliert, weshalb eine Rekodierung dieses Items erfolgte. *Tabelle 16* gibt einen Eindruck über die Formulierung der eingesetzten Items.

Tabelle 16. *Beispielitem zur Skala "Lehrerwahrnehmung"*

Skala	Beispielitem
Lehrerwahrnehmung	Unser Mathematiklehrer / Unsere Mathematiklehrerin unterrichtet mit Begeisterung.

In Anlehnung an die Kennwerte des vorangegangenen Abschnittes werden auch hier die Reliabilitätsmaße zu diesem Instrument berichtet (vgl. *Tabelle 17*), womit sich eine Einschätzung der Güte der Messung vornehmen lässt.

Tabelle 17. *Reliabilitätsmaße der Skala „Lehrerwahrnehmung“ (Schülerinnen und Schüler)*

Skala	Studie 1			Studie 2	
	Cronbachs Alpha / ICC(1) / ICC(2)			Cronbachs Alpha/ICC(1)/ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
Lehrerwahrnehmung	.88/ .34/ .92	.90/ .36/ .92	.90/ .34/ .92	.87/ .38/ .93	.89/ .41/ .93

Anmerkung.

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

Es zeigt sich, dass alle Reliabilitätsmaße zufriedenstellende bis gute Werte aufweisen.

5.3.6 Der standard-basierte Mathematiktest

Zu jedem Messzeitpunkt wurde ein Mathematiktest aus einem Pool von insgesamt 144 Items zusammengestellt, den die Schülerinnen und Schüler zu bearbeiten hatten.

5.3.6.1 Aufgabenauswahl

Zur Testung der mathematischen Kompetenzen der Schülerinnen und Schüler wurde in jeder Datenerhebung ein Mathematiktest eingesetzt, der zwischen 38 und 42 Teilaufgaben

bzw. Items beinhaltete. Der Test beinhaltet Aufgabenblöcke aus allen fünf mathematischen Leitideen, zu denen die Aufgaben jeweils zusammengefasst werden. Alle eingesetzten Items wurden in einer repräsentativen Normierungsstudie 2006 an mehr als 10.000 Neuntklässlern erprobt. Die so erlangten Schülerlösungen zu den Items führten durch die Kalibrierung zu Itemschwierigkeiten, die auf der kontinuierlichen Kompetenzskala der Bildungsstandards verankert wurden. Alle normierten Items werden in Blöcken zusammengefasst, die inhaltlich den mathematischen Leitideen zugeordnet werden können und sich an einer Bearbeitungszeit von insgesamt 20 Minuten pro Block orientieren. Innerhalb der Aufgabenblöcke sind die Items nach ihrer Schwierigkeit aufsteigend angeordnet. Damit wird eine Testzusammenstellung gewährleistet, die je nach Anlage und Untersuchungsplan eine variable Fokussierung auf das Testziel ermöglicht.

In der Begleitforschung zur Implementation der Bildungsstandards wurden Testhefte zusammengestellt, die jeweils vier der fünf möglichen Aufgabenblöcke (der Leitideen) enthielten und die Bearbeitungszeit auf ca. 80 Minuten festsetzten. Um Erinnerungseffekte zu minimieren, wurden in Studie 1 insgesamt zwei Mathematiktests eingesetzt. In Studie 2 wurde der Test zu beiden Testzeitpunkten nicht verändert. In allen Datenerhebungen wurden in jeder Klasse vier Versionen der Testhefte eingesetzt, deren Verteilung auf die Schülerinnen und Schüler zufällig erfolgte. Jeder Schüler bearbeitete somit in jedem Test Items aus vier Leitideen.

5.3.6.2 Testdesign

Die vier Versionen der eingesetzten Testhefte eines jeden Tests wurden miteinander verlinkt. Dazu wurde der Aufgabenblock aus der Leitidee 1 „Zahl“ in jedes Testheft eingebaut. Aufgrund der geltenden Globalskala des Kompetenzmodells Mathematik werden die Itemschwierigkeiten und Personenfähigkeiten zusammen skaliert, d.h. sie werden nicht nach Leitideen getrennt betrachtet, wodurch die Verlinkung über einen speziellen Aufgabenblock problemlos erfolgen kann. Abbildung 15 gibt einen Überblick über die eingesetzten Tests und deren Testdesign.

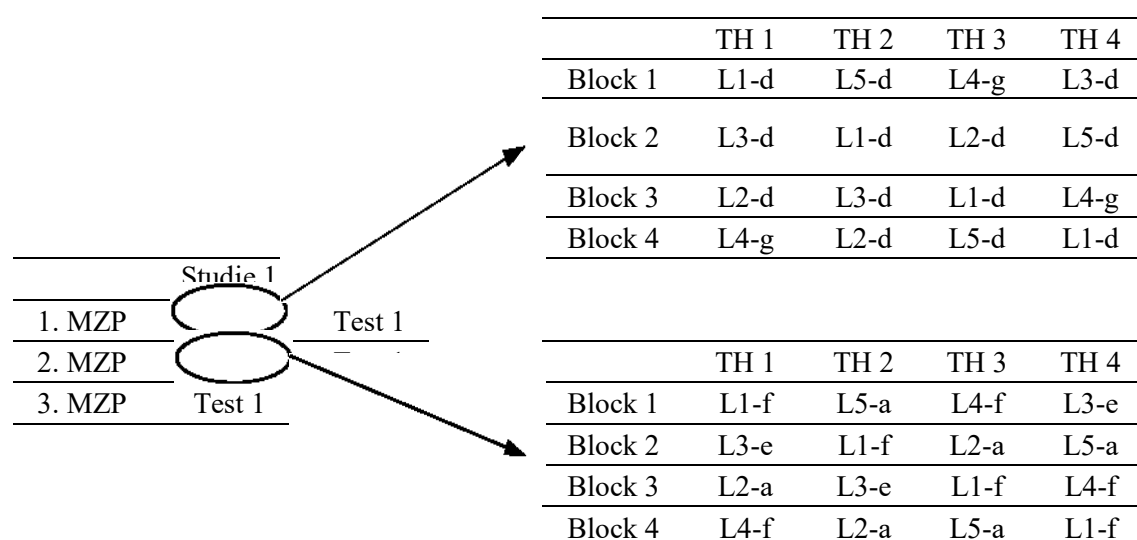


Abbildung 15. Design des Mathematiktests.

Die Testhefte des Tests 1 variieren in ihrer Schwierigkeit um 11 Kompetenzpunkte.

Testheft 1 ist mit einer Schwierigkeit von 525 Punkten auf der Berichtsmetrik der Bildungsstandards das schwerste Testheft; Testheft 3 wird mit einer gemittelten Itemschwierigkeit von 514 Punkten als leichtestes Testheft innerhalb des Tests 1 identifiziert. Alle eingesetzten Testhefte sind so konzipiert, dass sie die Schülerinnen und Schüler auf das Erreichen des Regelstandards testen.

Der Test 2 (zweiter Messzeitpunkt) weist in seiner Schwierigkeit eine Breite von 484 Punkten (Testheft 4) bis 501 Punkten (Testheft 1) auf. Insgesamt ist der Test 2 um 25 Kompetenzpunkte leichter als der Test 1. Folglich ist auch der Test 2 auf der Kompetenzstufe 3 einzuordnen.

Jeweils nach der Bearbeitung des zweiten Aufgabenblockes wurde eine Pause eingeplant, die je nach Schulsituation zwischen 5 und 15 Minuten dauerte. Danach erfolgte die Bearbeitung von zwei weiteren Aufgabenblöcken. Wenngleich innerhalb eines jeden Tests eine Verlinkung der verschiedenen Testheftversionen stattfand, so waren die Tests 1 und 2 nicht miteinander verlinkt, d.h. im Test 2 wurden andere Aufgaben eingesetzt als in Test 1.

Alle Testungen wurden von geschulten Testleitern administriert. Die Dateneingabe erfolgte zum ersten Messzeitpunkt der Studie 1 manuell, alle anderen Testungen wurden computerbasiert eingelesen. Die halboffenen und offenen Aufgabenstellungen wurden von einem geschulten Kodierer nach einem Richtig/Falsch-Schema bewertet. Tabelle 18 gibt einen Überblick über die Gütekriterien der Tests.

Tabelle 18. Gütekriterien zum Mathematiktest der Schülerinnen und Schüler

Skala	Studie 1			Studie 2	
	Cronbachs Alpha* / ICC(1) / ICC(2)			Cronbachs Alpha* / ICC(1) / ICC(2)	
	MZP 1	MZP 2	MZP 3	MZP 1	MZP 2
Mathematik-test	.85/ .52/ .96	.88/ .49/ .95	.85/ .49/ .95	.87/ .53/ .96	.89/ .54/ .96

Anmerkung.

* EAP/PV Reliabilität aus ConQuest, vergleichbar mit Cronbachs Alpha

ICC(1) = Intraklassenkorrelation berechnet als $ICC(1) = \tau^2 / (\tau^2 + \sigma^2)$

ICC(2) = Reliabilität der Klassenmittelwerte berechnet als $ICC(2) = (k * ICC(1)) / [1 + (k - 1) * ICC(1)]$

Die Reliabilitätskennzahlen weisen akzeptable Werte auf. Die Intraklassenkorrelationen ICC(1) und ICC(2) weisen auf eine hohe Güte der eingesetzten Tests hin.

5.4 Vorgehensweise und Analyse von Längsschnittdaten

Die Arbeit mit Längsschnittdaten bedarf einer vorherigen intensiven Datenaufbereitung. Für die zu untersuchenden Forschungsfragen lagen die Daten jeweils getrennt nach Erhebungszeitpunkt und Ort der Erhebung vor. Über eine Personenidentifikationsnummer, die die Teilnehmer zu jedem Messzeitpunkt persönlich aus nahezu unveränderlichen Daten erzeugen konnten, gelang es, die Erhebungswellen zu einem Längsschnittdatensatz zusammenzufügen. Die Personenidentifikationsnummer setzte sich dabei folgendermaßen zusammen:

1. Letzter Buchstabe des Vornamens des Vaters
2. Erster Buchstabe des Vornamens der Mutter
3. Die erste Zahl der Hausnummer
4. Die letzte Zahl der Telefonnummer
5. und 6. Die Ziffern des Geburtstages der Mutter.

Bei Abweichungen im so genannten „Persönlichen Code“ erfolgte ein Vergleich der Identifikationsnummern über weitere Variablen wie Schulname, Klassenname und Geschlecht. Bei erkläraren Abweichungen von nur einer Ziffer (Umzug, geänderte Telefonnummer), erfolgte dennoch ein Matching der Datensätze. Die weitere Vorgehensweise nach Generierung der Längsschnittdatensätze wird im Folgenden erläutert.

5.4.1 Generierung eines Zeitfaktors

In Studie 1 wurde eine Zeitvariable eingefügt, die den Zeitraum von der ersten Messung der ersten Schule bis zum Messzeitpunkt der jeweils interessierenden Schule dokumentiert. Die Zeitvariable hat für jede Schule drei Ausprägungen. Schulen, die im Februar 2007 getestet wurden, erhalten als erste Ausprägung der Zeitvariablen eine Null. Schulen, die aus organisatorischen Gründen erst im März getestet wurden, schreiben als ersten Zeitwert eine eins, da der Testtermin einen Monat nach den ersten getesteten Schulen realisiert werden konnte und dieser Lernfortschritt damit in Rechnung gestellt wird. Für die zweite Ausprägung jeder Schule werden die Monate in Rechnung gestellt, die seit der Testung der ersten Schule (und damit seit Beginn der Intervention) verstrichen sind. Ebenso wurde die Zeitausprägung zum dritten Messzeitpunkt ermittelt. Die Zeitvariable zeigt zum ersten Messzeitpunkt damit eine Ausprägung von Null oder Eins, womit impliziert wird, dass alle Schulen entweder im Februar 2007 oder März 2007 getestet wurden. Zum zweiten Messzeitpunkt wird den Schulen entweder eine acht (acht Monate nach Februar 2007), eine neun oder eine zehn zugeteilt. Der Testzeitraum zum zweiten Messzeitpunkt erstreckt sich folglich von Oktober bis Dezember 2007. Entsprechend wurde mit dem dritten Messzeitpunkt verfahren, der sich in der Zeit von März bis April 2008 befand und zu den Zeitangaben dreizehn oder vierzehn seit Beginn der Intervention führte (vgl. Anhang B).

Die Skalierung der Zeit-Variable in Monaten ist im Hinblick auf die Interpretation der Ergebnisse insofern entscheidend, als dass sie die Maßeinheit der Wachstumseffekte widerspiegelt. Ein auftretender Effekt ist damit als monatliches Wachstum in Höhe des Parameters zu interpretieren. In den Modellgleichungen zur Beantwortung der Hypothesen wird die Zeitvariable als Prädiktor in das Modell auf der Ebene 1 aufgenommen

$$Y_{it} = \pi_{0i} + \pi_{1i}a_{it} + \varepsilon_{it}$$

wobei a_{it} den Messzeitpunkt in Monaten ab erster Messung für Person i beschreibt.

5.4.2 Skalenbildung

Alle Items aus den in Abschnitt 5.3 vorgestellten Instrumenten wurden mit Ausnahme des Leistungstests zu Mittelwertskalen zusammengefasst. In Auswertung der vorliegenden Reliabilitäten für die Items eines Merkmals wurden alle eingesetzten Items verwendet. Eine Ausnahme stellt der erstmals eingesetzte Fragebogen zu den kompetenzorientierten Schülertätigkeiten dar. Hier wurden für jede Kompetenz mehr als fünf Items entwickelt.

Als Ziel wurde jedoch von Anfang an die einheitliche Erfassung einer Kompetenz mit fünf Items betrachtet.

5.4.3 Skalierung von Personenfähigkeiten

Zur Ermittlung von Personenfähigkeiten im standard-basierten Mathematiktest wurde der rekodierte und somit in einem 0-1-Format vorliegende Datensatz zur Skalierung an das Programm ConQuest (Wu, Adams & Wilson, 1998, Version 2.0) übergeben. Die interessierenden Personenfähigkeiten wurden auf der Basis des Rasch-Modells generiert, das es erlaubt, Itemschwierigkeiten und Personenfähigkeiten auf einer gemeinsamen Skala abzubilden. Die aus der Normierung bekannten Itemschwierigkeiten auf der Logit-Skala dienten dabei als fixierte Ankerparameter. Damit wird sichergestellt, dass die Berechnung der Personenfähigkeiten unter der Voraussetzung erfolgt, dass die Items zu jedem Messzeitpunkt die gleiche Schwierigkeit besitzen⁴⁸. Mit diesem Vorgehen lassen sich die ermittelten Personenparameter als Werte auf der derselben Skala interpretieren und eine Differenz der Fähigkeitswerte zwischen zwei Messzeitpunkten kann als Fähigkeitszuwachs interpretiert werden (Hartig & Kühnbach, 2006). Die hier ermittelten Personenfähigkeiten wurden als Weighted Likelihood-Schätzungen (WLE) aus dem Rasch-Modell berechnet. Diesem Maximum-Likelihood-Schätzer werden geringe Verzerrungen und eine hohe Erwartungstreue zugeschrieben (Walter, 2005). Bei der Berechnung des WLE als Schätzer für die Personenfähigkeit im Rahmen von IRT-Modellen wird die Wahrscheinlichkeit maximiert, das beobachtete Lösungsmuster einer Person unter Berücksichtigung der bearbeiteten Items zu zeigen. Trotz unvermeidlicher Messfehler bietet die Verwendung von WLEs als Fähigkeitsschätzer eine Korrektur der Streuung, die bei der Analysen auf Basis von Maximum-Likelihood-Schätzern gewöhnlich überschätzt wird (Hartig & Kühnbach, 2006). Die Personenfähigkeiten werden als sogenannte Logits ausgegeben, die sich auf einer kontinuierlichen Fähigkeitsskala verteilen und mit den Itemschwierigkeiten kooperieren.

In der Kommunikation um die Bildungsstandards wird eine Metrik verwendet, die den Mittelwert dieser Personenfähigkeiten durch eine Transformation auf 500 setzt und die Standardabweichung auf 100. Diese Transformation erfolgt durch die Gleichung,

$$BISTA - Metrik = 500 + \text{Logit} \cdot \frac{100}{1.105}$$

⁴⁸ Damit wird folglich außer Acht gelassen, dass die Items für eine bestimmte Schülergruppe schwieriger sein könnten.

wobei 1.105 die Standardabweichung der neunten Jahrgangsstufe in der Normierung der Bildungsstandards aus den Jahren 2006 und 2007 darstellt. Die so erhaltene Personenfähigkeit eines jeden Schülers lässt sich dann in das Kompetenzstufenmodell einordnen und um die für die jeweilige Stufe geltende Kompetenzzuschreibung ergänzen.

5.4.4 Umgang mit fehlenden Werten

Im Umgang mit Längsschnittstudien stellt sich gewöhnlich die Problematik des Umgangs mit fehlenden Werten weit verschärfter dar als in Querschnittstudien, weshalb die Behandlung der Thematik bei der Analyse solcher Datensätze wie auch den hier vorliegenden unumgänglich ist. Während in querschnittlichen Designs das Vorkommen von fehlenden Werten auf einzelnen Variablen oder ganze Skalen innerhalb eines Fragebogens begrenzt bleibt, erweitert sich die Problematik in Längsschnittstudien auf das Fehlen ganzer Fälle für bestimmte Erhebungswellen⁴⁹. Weiterhin beeinflussen diese fehlenden Werte im Längsschnitt nicht nur die Daten in einer speziellen Erhebungswelle, sondern gleichzeitig auch die Daten dieser Personen über den gesamten Erhebungszeitraum. Das wirkt sich auf die Stichprobengröße aus und wirkt mitunter auf die zur Verfügung stehenden Analysemethoden. Menard (2002) führt verschiedene Handlungsweisen für die drei Arten von Missing Values auf und weist gleichzeitig darauf hin, dass keine der aufgezeigten Handlungsempfehlungen eine komplett zufriedenstellende Lösung darstellt. In der Theorie zur Analyse von fehlenden Werten werden nach der Terminologie von Rubin (1976) grundsätzlich drei Arten solcher Datenausfälle unterschieden: (1) *missing completely at random* (MCAR), wobei das Fehlen von Daten ausnahmslos als zufällig angesehen wird, d.h. das Fehlen eines Wertes kovariiert nicht mit der Ausprägung anderer Variablen und hängt nicht mit den Werten der kritischen Variablen zusammen, (2) fehlende Daten werden als *missing at random* (MAR) eingeordnet, wenn das Fehlen in Abhängigkeit von einer manifesten Variable auftritt, aber unabhängig von einer latenten Variablen. Damit unterscheidet sich die Bedeutung des Wortes „random“ von dem Verständnis des Begriffes im psychologischen Bereich in dem Sinne, dass es sich um ein „bedingtes Fehlen“ des Wertes handelt. Anders ausgedrückt, nach Kontrolle der beobachteten Variablen hängt das Fehlen eines Wertes nicht mehr von der Ausprägung der latenten Variablen ab – das Fehlen des Wertes ist „zufällig“.

⁴⁹ Bijleveld & van der Kamp (1998, S. 41) führen dieser Unterteilung noch die missing subjects dazu und definieren diese, als Personen, die zur Stichprobe gehören, jedoch nie an der Befragung teilnahmen, die so genannten Non-Responses.

Die letzte Form der fehlenden Werte wird mit (3) *missing not at random* (MNAR) bezeichnet und tritt ein, wenn das Fehlen von Daten weder „vollständig zufällig“ noch „zufällig“ auftritt. Damit wird das Fehlen „nicht zufällig“ und bedeutet, dass auch nach Kontrolle einer beobachteten Variablen das Fehlen von Werten von den Ausprägungen einer anderen Variablen abhängt. Diese Art der fehlenden Werte sind die am häufigsten auftretenden und am schwierigsten im Umgang. Das Problem mit MNAR ist, dass die fehlenden Werte jeden möglichen Wert annehmen könnten. Graham (2009) führt dazu aus, dass die fehlenden Werte vom MAR in Längsschnittstudien durch die Werte der anderen Wellen vorhergesagt werden können, wobei er einräumt, dass Missings nie in totaler Reinform der Kategorien MCAR, MAR oder MNAR auftreten. Die Unsicherheit bei der Bestimmung der Art der fehlenden Werte wird reduziert, wenn mehr als zwei Erhebungswellen Daten liefern und damit ein Trend sichtbar wird.

Die Lösungen zum Umgang mit fehlenden Werten ergeben sich in der Regel zweckgebunden. Bei großen Stichproben und kleinen Ausfallraten ergibt sich als effizienteste Lösung oftmals die Nichtbeachtung der Personenantworten für die spezifische Modellberechnung (listwise oder pairwise deletion)⁵⁰. Eine andere relativ einfache und effektive Handhabung wird in dem Ersatz des fehlenden Items durch den Skalenmittelwert gesehen oder aber in der Gewichtung verschiedener Fälle zum Ausgleich beim Ausfall bedingter Items, die mit anderen verlinkt sind (z.B. sozioökonomischer Status). Für die komplexeren Problematiken des Fehlens ganzer Skalen oder Fälle in einzelnen Erhebungswellen bieten sich die Verfahren der multiplen Imputation an (Menard, 2002). Graham (2009, S. 562) empfiehlt als Verfahren im Umgang mit fehlenden Werten in Längsschnittstudien „...any missing data procedure that preserves (i.e., estimates without bias) variances, covariances, and means is acceptable.“.

5.4.5 Datenimputation

In Studie 1 ist für die untersuchten Merkmale ein ansteigender Missinganteil zu beobachten, der auf die sinkende Teilnahmebereitschaft zurückzuführen ist. Für den zugrunde liegenden Datensatz wurden die einzelnen Messzeitpunkte zusammengefügt und über den in jeder Erhebung abgefragten „persönlichen Code“ verbunden. So ergibt sich ein Gesamtdatensatz mit 152 Teilnehmern und Teilnehmerinnen. Tabelle 19 stellt den Missinganteil für die Stichprobe der Lehrkräfte in Studie 1 dar und deutet auf eine hohe Fluk-

⁵⁰ Menard (2002) empfiehlt die Anwendung des listenweisen Fallausschlusses mit Bezug zu Graham und Hofer (2000) nur bei einem sehr kleinen Fallverlust von weniger als 5 Prozent.

tuation bei der Teilnahme an den einzelnen Untersuchungen hin. Zu beachten ist hierbei jedoch, dass jede Person, die zu einem späteren Zeitpunkt der Studie beitrug, die Stichprobe und damit die Referenzanzahl an Lehrkräften erhöhte. So erklären sich auch die Ausfallquoten zum ersten Messzeitpunkt⁵¹.

Tabelle 19. Missinganteile in Prozent in der Studie 1 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Lehrkräfte

Skala	MZP 1	MZP 2	MZP 3	MCAR Test nach Little
Fachkonferenz	27.0	46.1	52.6	$\chi^2=11.31; p = .255$
Kooperation	27.0	46.7	52.0	$\chi^2=20.57; p = .015$
Einstellung positiv	26.3	49.3	54.6	$\chi^2=7.41; p = .595$
Einstellung negativ	26.3	49.3	54.6	$\chi^2=21.79; p = .010$
math. Argumentieren	26.3	46.1	52.0	$\chi^2=11.44; p = .247$
Probleme math. Lösen	27.0	46.1	52.0	$\chi^2=4.17; p = .900$
math. Modellieren	27.0	46.1	52.0	$\chi^2=11.00; p = .275$
Darstellungen Verwenden	27.0	46.1	52.0	$\chi^2=5.03; p = .831$
technisch Arbeiten	27.0	46.1	52.0	$\chi^2=15.59; p = .076$
math. Kommunizieren	26.3	46.1	52.0	$\chi^2=3.45; p = .944$

Anmerkung. Analyse der fehlenden Werte auf Skalenebene; als Referenz dient der Gesamtdatensatz mit N = 152.

Im Messzeitpunkt 3 sind für die Lehrkräfte die höchsten Ausfallquoten zu beobachten. Dennoch beantworteten 44 Lehrkräfte den Fragebogen zu allen drei Erhebungszeitpunkten und 72 Personen zu mindestens zwei Erhebungen. Zur Aufrechterhaltung eines möglichst vollständigen Längsschnittdatensatzes wird folgendes Kriterium im Hinblick auf die Behandlung fehlender Werte festgelegt:

- Personen mit nur einem gültigen Messzeitpunkt fallen aus der Betrachtung des Längsschnittes heraus,
- Personen mit mindestens zwei gültigen Erhebungen in unbestimmter Teilnahme-kombination verbleiben im Datensatz; die fehlenden Werte werden ersetzt.

Da die Teilnahme der Schülerinnen und Schüler zu einem großen Teil von den Lehrkräften gesteuert wird, ergibt sich hier ein anderes Bild. Der relativ hohe Ausfall zum zweiten

⁵¹ Nehmen bspw. in der zweiten Erhebung 10 Personen zusätzlich zu den 100 Personen aus der ersten Erhebung teil, dann erhöht sich die Referenzgröße zur Berechnung der fehlenden Werte auf 110. Damit ergibt sich auch für den ersten Messzeitpunkt eine Ausfallquote von 9.1 Prozent.

Messzeitpunkt ist auf den Wegfall ganzer Klassen zurückzuführen, die entweder aus der Studie ausgestiegen waren oder aus organisatorischen Gründen keine Erhebung ermöglichen konnten (Tabelle 20).

Tabelle 20. *Missinganteile in Prozent in der Studie 1 nach Untersuchungsmerkmal und Erhebungszeitpunkt, Schüler*

Skala	MZP 1	MZP 2	MZP 3	MCAR Test nach Little
Selbstkonzept	26.7	40.3	29.2	$\chi^2=24.71$; $p = .003$
Interesse	26.7	40.3	29.3	$\chi^2=5.29$; $p = .809$
Lehrerwahrnehmung	27.3	40.9	29.3	$\chi^2=14.37$; $p = .110$
math. Argumentieren	27.4	40.4	29.2	$\chi^2=51.93$; $p = .000$
Probleme math. Lösen	27.5	40.4	29.2	$\chi^2=36.96$; $p = .000$
math. Modellieren	27.5	40.5	29.1	$\chi^2=13.72$; $p = .133$
Darstellungen Verwenden	27.5	40.4	29.2	$\chi^2=31.74$; $p = .000$
technisch Arbeiten	27.4	40.4	29.2	$\chi^2=82.73$; $p = .000$
math. Kommunizieren	27.4	40.4	29.1	$\chi^2=37.86$; $p = .000$
Mathematikleistung	26.5	39.7	28.0	$\chi^2=248.48$; $p = .000$

Anmerkung. Analyse der fehlenden Werte auf Skalenebene; als Referenz dient der Gesamtdatensatz mit $N = 1835$.

Während der *missing completely at random* (MCAR) – Test nach Little zur Untersuchung der Art der fehlenden Werte für die Lehrkräfte nahezu vollständig auf MCAR hinweist (mit Ausnahme der Merkmale Kooperation und negative Einstellungen gegenüber den Bildungsstandards), wird innerhalb der Schülerstichprobe MCAR mehrheitlich abgelehnt. Es zeigt sich, dass ca. 780 Schülerinnen und Schüler an allen drei Messzeitpunkten und ca. 1.247 Lernerinnen und Lerner an mindestens zwei Erhebungen teilnahmen. Zur Vermeidung eines starken Datenverlustes wird auch in der Schülerstichprobe der Studie 1 das für die Lehrkräfte geltende Kriterium der mindestens zweimaligen Testung angewandt, um die längsschnittliche Entwicklung abzubilden.

Aufgrund der auftretenden Ablehnung der MCAR-Hypothese werden die fehlenden Werte mit dem Verfahren der multiplen Datenimputation ersetzt. Orientiert an der gängigen Literatur (Kenward & Carpenter, 2009, Rubin, 1987) zum Umgang mit fehlenden Werten werden die Datensätze der Lehrkräfte und Schüler um fünf Imputationen erweitert⁵².

⁵² Einstellungen in SPSS: Startwert Zufallsgenerator auf 2000000, Full Conditional Model (FCM).

Mit Ausnahme des Mathematikleistungswertes für die Schülerinnen und Schüler erfolgt die multiple Imputation der fehlenden Werte auf Itemebene⁵³. Für die Schätzung der fehlenden Werte werden nur die Items der dazugehörigen Skala als Prädiktor verwendet. Im Anschluss an das Imputationsverfahren wurde die Skalenbildung durchgeführt. Für die Berechnungen wurde die Software SPSS verwendet, deren *fully conditional specification*-Methode (FCS) für jede Variable mit fehlenden Werten ein Imputationsmodell spezifiziert, das in einem iterativen Prozess die fehlenden Werte hier fünf Mal ersetzt.

In Studie 2 ergibt sich durch das Zusammenfügen der Einzeldateien der Messzeitpunkte ein Gesamtdatensatz von $N=222$ Lehrkräften. Davon nahmen 171 Personen an der ersten Erhebung teil und 135 Personen an der zweiten Erhebung. Es konnten 84 Lehrkräfte identifiziert werden, die an beiden Messzeitpunkten gültige Daten aufweisen. Tabelle 21 bildet die Missinganteile in Prozent zum jeweiligen Messzeitpunkt für die einzelnen Skalen ab. Dieses geänderte Vorgehen im Vergleich zur Studie 1 liegt in dem Imputationsdesign der Studie 2 begründet.

Tabelle 21. *Missinganteile in Prozent in der Studie 2 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Lehrkräfte*

Skala	N_1	MZP 1	N_2	MZP 2
Fachkonferenz	166	2.9	133	1.5
Kooperation	165	3.5	134	0.7
Einstellung positiv	164	4.1	130	3.7
Einstellung negativ	165	3.5	130	3.7
math. Argumentieren	170	0.6	134	0.7
Probleme math. Lösen	170	0.6	134	0.7
math. Modellieren	170	0.6	134	0.7
Darstellungen Verwenden	170	0.6	134	0.7
technisch Arbeiten	170	0.6	134	0.7
math. Kommunizieren	170	0.6	134	0.7

Anmerkung. Als Referenz im Querschnitt: $N_1 = 171$ und $N_2 = 135$.

Durch das Design der Studie 2 mit nur zwei Erhebungswellen kann das oben beschriebene Kriterium zur multiplen Datenimputation keine Anwendung mehr finden, da von Personen mit einem fehlenden Messzeitpunkt nur *ein* weiterer Referenzwert vorliegt. Daher

⁵³ Die Leistungswerte der Schülerinnen und Schüler wurden als Skalen- bzw. Testwert nach den angegebenen Kriterien ersetzt.

erfolgt eine multiple Imputation der fehlenden Werte im Querschnitt. Erst im Anschluss an die Imputation erfolgt hier das Matching der Datensätze und die Skalenbildung.

Die gleiche Vorgehensweise wurde auch für die Schülerstichprobe der Studie 2 angewandt. Hier erfolgte ebenso eine multiple Imputation der fehlenden Werte pro Querschnitt.

Tabelle 22. *Missinganteile in Prozent in der Studie 2 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Schüler*

Skala	N ₁	MZP 1	N ₂	MZP 2
Selbstkonzept	2373	3.5	2092	2.7
Interesse	2371	3.5	2089	2.9
Lehrerwahrnehmung	2356	4.1	2079	3.3
math. Argumentieren	2353	4.3	2086	3.0
Probleme math. Lösen	2353	4.3	2085	3.1
math. Modellieren	2351	4.4	2083	3.2
Darstellungen Verwenden	2355	4.2	2086	3.0
technisch Arbeiten	2354	4.2	2087	3.0
math. Kommunizieren	2354	4.2	2085	3.1
Mathematikleistung	2458	0.0	2151	0.0

Anmerkung. Als Referenz im Querschnitt: $N_1 = 2458$ und $N_2 = 2151$.

5.4.6 Mehrebenenstruktur

In Anlehnung an Singer und Willet (2003) wurde der vorliegende Längsschnittdatensatz der Lehrkräfte in einen „Person-Period-Datensatz“ bzw. in ein univariates Format umgewandelt (so genanntes Long-Format), womit die Messzeitpunkte jeder Lehrkraft nicht als einzelne Variablen erscheinen, vielmehr bleibt die wiederholt gemessene Variable fix und die Werte der einzelnen Erhebungswellen werden als neuer Fall aufgeführt. Damit wird die Anzahl der Variablen verringert, die Anzahl der Fälle erhöht sich jedoch. Die Verbindung der Messwerte einer Person zu verschiedenen Zeitpunkten erfolgt durch die Personenkennziffer (ID). Hinsichtlich der interessierenden Outputvariablen werden Skalenwerte als Summenscores der Items gebildet (Abschnitt 5.4.2), womit sich die untersuchten Outputvariablen folgendermaßen darstellen:

- *Einstellung gegenüber der Einführung von Bildungsstandards*
 - positive Aspekte bei der Einführung von Bildungsstandards
 - negative Aspekte bei der Einführung von Bildungsstandards

- *Kooperation in der Fachkonferenz*
- *Zusammenarbeit in der Fachkonferenz*
- *kompetenzorientierte Unterrichtsgestaltung:*
 - mathematisch Argumentieren
 - mathematische Probleme Lösen
 - mathematisch Modellieren
 - Darstellungen verwenden
 - technisch Arbeiten
 - mathematisch Kommunizieren.

In einem ersten Schritt wird daher ein unkonditioniertes Modell ohne erklärende Prädiktoren auf beiden Ebenen gerechnet, womit sich der Intra-Klassen-Korrelationskoeffizient (ICC) bestimmen lässt (Modell A).

In einem zweiten Schritt wird ein unkonditioniertes Wachstumsmodell spezifiziert (Singer und Willett, 2003), das auf der Ebene 1 einen Wachstumsparameter quantifiziert (Modell B). Dieses Modell repräsentiert das individuelle Wachstum jeder Person, wobei jede Abweichung der beobachteten Daten von der linearen Wachstumsfunktion auf einen zufälligen Messfehler zurückgeführt wird. Für beide Studien besteht aufgrund der geringen Anzahl an Messzeitpunkten die Annahme eines linearen Wachstums⁵⁴.

Die grundsätzlichen Hypothesen über die Entwicklung in den Untersuchungsmerkmalen für beide Gruppen werden mit einem Modell beantwortet, das den zeitinvarianten Faktor „Gruppe“ in die Regressionen miteinbezieht. Dabei entspricht das Modell auf Ebene 1 dem unkonditionierten Wachstumsmodell, auf Ebene 2 wird zusätzlich eine Gruppenvariable für beide Koeffizienten aufgenommen, womit nicht nur der Ausgangswert beider Gruppen unterschieden werden kann, sondern ebenso das Wachstum nach Gruppen getrennt analysiert wird (Modell C). Die Modellgleichung für die Ebene 1 stellt sich folgendermaßen dar:

$$Y_{ti} = \pi_{0i} + \pi_{1i}a_{ti} + \varepsilon_{ti}$$

Auf der Ebene 2 ergibt die Aufnahme der Gruppenvariablen für Studie 1 folgende Gleichungen:

$$\pi_{0i} = \beta_{00} + \beta_{01}X_{qi} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}X_{qi} + r_{1i}$$

Die Variable X stellt dabei die Gruppenvariable dar, die dichotom in den Ausprägungen Null für Vergleichsschulen und Eins für Projektschulen definiert ist.

⁵⁴ Raudenbush und Bryk (2002) beziffern eine geringe Anzahl an Messzeitpunkten pro Person mit drei bis vier.

In Studie 2 mit nur zwei Messzeitpunkten wird die Varianz des Wachstumsfaktors auf der Ebene 2 auf Null gesetzt. Dieses Vorgehen war nötig, um die Anzahl der Freiheitsgrade in der Modellgleichung zu reduzieren (Modell D).

Das Modell auf der Ebene 2 stellt sich wie folgt dar:

$$\begin{aligned}\pi_{0i} &= \beta_{00} + \beta_{01} * X_{qi} + r_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11} X_{qi}\end{aligned}$$

Die Analyse der Wachstumsmodelle erfolgte mit der Software HLM (Raudenbush & Bryk, 2002). Als Schätzmethode wurde die Restricted Maximum Likelihood (RML) gewählt.

6 Ergebnisse

6.1 Drop-Out Analysen

In Anlehnung an den Abschnitt 5.4.4 und wie für Längsschnittanalysen üblich, werden die Schwankungen in den Teilnehmerzahlen auf eine Systematik hin untersucht. Hierbei liegt der Fokus auf der Beschreibung zweier Gruppen von Personen in der Gesamtstichprobe. Die Gruppe der als *Abbrecher* bezeichneten Lehrkräfte nahm an der ersten Datenerhebung teil, jedoch nicht an der zweiten Befragung. Für die als *Teilnehmer* bezeichnete Gruppe von Personen gelten in der Studie 1 verschiedene Teilnahmevariationen, die jedoch alle die Teilnahme am Messzeitpunkt 1 erfordern und mindestens eine weitere Teilnahme aufweisen. Die nach dem ersten Messzeitpunkt nicht an der darauf folgenden Datenerhebung teilgenommen hat und andererseits die Gruppe (als Teilnehmer bezeichnet), die die Befragung und Testung weiter aktiv unterstützte. Das Charakteristikum mehrerer Messzeitpunkte in beiden Studien führt dazu, dass in den folgenden Analysen nur die Lehrkräfte und Schüler einfließen, die zum ersten Messzeitpunkt an der Befragung und Testung teilnehmen und zum zweiten Messzeitpunkt keine Daten einbrachten. Andere Teilnahmevariationen, wie etwa die verspätete Teilnahme an der Befragung oder die Kombination der Teilnahme zum ersten und dritten Messzeitpunkt (Studie 1) bleiben hier unberücksichtigt.

6.1.1 Drop-Out-Analyse der Studie 1

Zum ersten Messzeitpunkt nahmen insgesamt 113 Lehrkräfte an der Studie 1 teil, der zweite Messzeitpunkt umfasst 82 Lehrkräfte (Abschnitt 5.1.5). Von diesen 113 Lehrkräften füllte nur 72 den Lehrerfragebogen zum zweiten Messzeitpunkt aus. Differenziert man noch einmal nach Projekt- und Vergleichsschulen, dann zeigt sich, dass von den ins-

gesamt 63 Lehrkräften an den Projektschulen, mehr als die Hälfte die Mitarbeit zurücknahm (59 Prozent, Tabelle 23). Mehrheitlich waren es die männlichen Lehrkräfte, die ihre weitere Teilnahme versagten. Hinsichtlich der Schulform lässt sich keine eindeutige Systematik erkennen und in Bezug auf das Alter wird sichtbar, dass in jeder Altersgruppe tendenziell die Hälfte der Lehrkräfte das Projekt fortsetzen. Lediglich in der Altersgruppe der 20-39-jährigen brechen mehr als 50 Prozent das Projekt ab. setzen sich die Abbrecher zur Hälfte aus Lehrkräften im Alter zwischen 40-49 Jahren zusammen. Es wird deutlich, dass die Gruppe der Abbrecher sich gegenüber der Gruppe Teilnehmer besonders im Merkmal „Gruppe“ unterscheidet. In der Gruppe Abbrecher können besonders viele Personen identifiziert werden, die in den Projektschulen arbeiten. Für die anderen Merkmale wird eine so starke Auffälligkeit nicht beobachtet.

Tabelle 23. Deskriptive Statistiken (in Prozent) für die Lehrergruppen Teilnehmer und Abbrecher in Studie 1, Lehrkräfte

Merkmal		Abbrecher (n=80)	Teilnehmer (n=72)	χ^2	df	p
Gruppe	Projektschulen (n=63)	58.7	41.3	1.61	1	.205
	Vergleichsschulen (n=89)	48.3	51.7			
Bundesland	Berlin (n=76)	59.2	40.8	2.64	1	.104
	Brandenburg (n=76)	46.1	53.9			
Geschlecht	männlich (n=54)	55.6	44.4	0.81	1	.367
	weiblich (n=92)	47.8	52.2			
Alter	25 – 29 (n=1)	0.0	100.0	4.76	4	.313
	30 – 39 (n=13)	53.8	46.2			
	40 – 49 (n=55)	45.5	54.5			
	50 – 59 (n=39)	43.6	56.4			
	60 und älter (n=2)	50.0	50.0			
Schulform	Gymnasium (n=73)	53.4	46.6	0.04	1	.851
	Andere (n=79)	51.9	48.1			

Es wird deutlich, dass die Gruppe der Abbrecher sich gegenüber der Gruppe Teilnehmer besonders im Merkmal „Gruppe“ unterscheidet. In der Gruppe Abbrecher können besonders viele Personen identifiziert werden, die in den Projektschulen arbeiten. Für die anderen Merkmale wird eine so starke Auffälligkeit nicht beobachtet.

Durch eine Analyse der Mittelwertunterschiede für die interessierenden Untersuchungsmerkmale lassen sich die beiden Gruppen Abbrecher und Teilnehmer noch weiter unterscheiden. Hierzu wurde für die Teilnehmer zunächst ein Mittelwert für das untersuchte

Merkmal über die Messzeitpunkte gebildet. Für die Abbrecher ist per Definition nur ein Messwert zum untersuchten Merkmal vorhanden. Signifikante Unterschiede in den Mittelwerten lassen sich in den Merkmalen „Kooperation“ und „negative Einstellung gegenüber den Bildungsstandards“ feststellen. Mit Blick auf die Gruppenmittelwerte zeigt sich, dass die Gruppe der Abbrecher weniger stark kooperiert, den Bildungsstandards gegenüber negativer eingestellt ist als die Gruppe der Teilnehmer, in der Unterrichtsgestaltung die Schülerinnen und Schüler häufiger mathematisch argumentieren lässt, jedoch in geringerem Maße zu den Tätigkeiten des mathematischen Modellierens, des Verwendens von Darstellungen, des technischen Arbeitens und Kommunizierens auffordert. Nach Rost (2005) sind die signifikanten Mittelwertunterschiede in den Merkmalen technisch Arbeiten mit $\eta^2_p = .03$ als mittlere Effektstärke und in den Merkmalen Einstellungen negativ ($\eta^2_p = .11$) und Kooperation ($\eta^2_p = .10$) als große Effektstärken zu bewerten.

Tabelle 24. Mittelwertunterschiede zwischen den Abbrechern und Teilnehmern, Lehrerstichprobe

Merkmal	Gruppe	MW (SD)	F	p	η^2_p
Fachkonferenzarbeit	Teilnehmer (n=72)	2.49 (0.43)	0.44	.51	.003
	Abbrecher (n=78)	2.44 (0.55)			
Kooperation	Teilnehmer (n=72)	3.12 (0.48)	16.97	.00	.103
	Abbrecher (n=78)	2.78 (0.54)			
Einstellungen positiv	Teilnehmer (n=72)	3.52 (0.67)	3.55	.06	.024
	Abbrecher (n=75)	3.31 (0.68)			
Einstellungen negativ	Teilnehmer (n=72)	2.50 (0.51)	16.94	.00	.105
	Abbrecher (n=75)	2.90 (0.63)			
Argumentieren	Teilnehmer (n=72)	3.99 (0.70)	2.14	.15	.014
	Abbrecher (n=78)	4.16 (0.73)			
Probleme lösen	Teilnehmer (n=72)	4.28 (0.63)	0.01	.91	.000
	Abbrecher (n=77)	4.29 (0.78)			
Modellieren	Teilnehmer (n=72)	4.04 (0.74)	0.32	.57	.002
	Abbrecher (n=77)	3.96 (0.90)			
Darstellungen verwenden	Teilnehmer (n=72)	4.01 (0.64)	0.70	.40	.005
	Abbrecher (n=78)	3.91 (0.75)			
technisch Arbeiten	Teilnehmer (n=72)	4.67 (0.49)	5.49	.02	.036
	Abbrecher (n=77)	4.45 (0.64)			
Kommunizieren	Teilnehmer (n=72)	4.68 (0.66)	0.65	.42	.004
	Abbrecher (n=78)	4.59 (0.78)			

Da der Mathematikunterricht bei den Teilnehmern stärker durch das Abarbeiten von mathematischen Routinen (technisch Arbeiten) gekennzeichnet ist als bei den Abbrechern, könnte sich daraus das Interesse und der Wunsch zur weiteren Teilnahme am Projekt ableiten lassen.

Die Teilnahme der Schülerinnen und Schüler am Projekt hängt von der Zustimmung und Bereitschaft der Lehrkräfte ab. Dennoch zeigt sich in den Drop-Out-Analysen, wie die Schülerinnen und Schüler beschrieben werden können, die an der zweiten Testung nicht mehr teilnahmen. Als Teilnehmer werden entsprechend der Klassifizierung hier 1250 Schüler und Schülerinnen identifiziert. Von den Schülerinnen und Schülern der Projektschulen nehmen 73 Prozent weiterhin an der Studie teil, 27 Prozent brechen die Teilnahme ab.

Tabelle 25. Deskriptive Statistiken (in Prozent) für die Schülergruppe der Studie 1

Merkmal		Abbrecher (n=585)	Teilnehmer (n=1.250)	χ^2 / F	df	p
Gruppe	Projektschule	26.9	73.1	24.43	1	.000
	Vergleichsschulen	37.7	62.3			
Bundesland	Berlin	33.4	66.6	3.17	1	.075
	Brandenburg	29.4	70.6			
Geschlecht	männlich	34.2	65.8	16.21	1	.000
	weiblich	25.4	74.6			
Alter*	Geburtsjahr 89	66.1	33.9	105.43	1	.000
	Geburtsjahr 90	47.6	52.4			
	Geburtsjahr 91	25.4	74.6			
	Geburtsjahr 92	19.5	80.5			
	Geburtsjahr 93	14.8	85.2			
	Geburtsjahr 94	50.0	50.0			
	Missings	97.2	2.80			
Schulform	Gymnasium	16.0	84.00	162.63	1	.000
	Andere	44.0	56.00			

Anmerkung: *Für die Altersvariable wurde eine ANOVA durchgeführt. Es handelt sich hier um den F-Wert der Teststatistik.

In den Vergleichsschulen entscheiden sich 62 Prozent der Schülerinnen und Schüler für eine weitere Teilnahme an der Studie und 38 Prozent brachen nach der ersten Testung ab. Besonderes Interesse an dem Projekt zeigten die Gymnasiasten. Von diesen Schülerinnen und Schülern nahmen 84 Prozent auch an weiteren Testungen und Befragungen teil. Dieser Befund unterscheidet sich jedoch von der Teilnahmebereitschaft der gymnasialen Lehrkräfte. Es wird vermutet, dass hier ein großes Interesse an den Leistungsergebnissen der Schülerinnen und Schüler vorliegt, während die Lehrkräfte der Projektarbeit relativ verhalten gegenüberstehen (Tabelle 23).

Für die Schülerstichprobe der Studie 1 zeigen sich deutliche Unterschiede zwischen den Gruppen der Teilnehmer und Abbrecher (Tabelle 26). So weisen die Abbrecher ein gerin-

geres mathematisches Selbstkonzept auf und eine um mehr als 60 Punkte geringere Mathematikleistung. Ebenso können sie weniger stark einen kompetenzorientierten Unterricht wahrnehmen. Eine Ursache dafür könnte in der schlechteren Bewertung der Mathematiklehrkraft liegen. Die Abbrecher sind weniger zufrieden mit ihrer Lehrkraft als die Teilnehmer.

Tabelle 26. Mittelwertunterschiede zwischen Teilnehmern und Abbrechern in Studie 1, Schülerstichprobe

Merkmal	Gruppe	<i>MW (SD)</i>	<i>F</i>	<i>p</i>	η^2
mathematisches Selbstkonzept	Teilnehmer (n=1.248)	2.46 (0.70)	11.77	.001	.007
	Abbrecher (n=519)	2.33 (0.78)			
Interesse an Mathematik	Teilnehmer (n=1.248)	2.09 (0.64)	1.01	.315	.001
	Abbrecher (n=518)	2.13 (0.74)			
Lehrerwahrnehmung	Teilnehmer (n=1.247)	2.80 (0.72)	7.88	.005	.004
	Abbrecher (n=511)	2.69 (0.83)			
Argumentieren	Teilnehmer (n=1.246)	3.85 (0.87)	23.79	.000	.013
	Abbrecher (n=511)	3.61 (1.14)			
Probleme lösen	Teilnehmer (n=1.246)	3.73 (0.77)	22.71	.000	.013
	Abbrecher (n=510)	3.51 (1.05)			
Modellieren	Teilnehmer (n=1.247)	3.39 (0.90)	1.50	.222	.001
	Abbrecher (n=510)	3.33 (1.12)			
Darstellungen verwenden	Teilnehmer (n=1.246)	3.40 (0.77)	30.48	.000	.017
	Abbrecher (n=511)	3.16 (0.98)			
technisch Arbeiten	Teilnehmer (n=1.247)	3.98 (0.65)	59.90	.000	.033
	Abbrecher (n=511)	3.67 (0.95)			
Kommunizieren	Teilnehmer (n=1.247)	3.87 (0.84)	17.61	.000	.010
	Abbrecher (n=511)	3.67 (1.11)			
Mathematikleistung	Teilnehmer (n=1.249)	531.13 (95.15)	286.95	.000	.138
	Abbrecher (n=538)	447.39 (97.49)			

Mit Ausnahme der Merkmale „Interesse am Mathematikunterricht“ und „mathematisches Modellieren“ weisen alle Merkmale signifikante Gruppenunterschiede auf. Somit zeigen die Teilnehmer und Abbrecher ein gleich großes Interesse am Fach Mathematik und haben in etwa die gleiche Wahrnehmung zum mathematischen Modellieren. Für alle anderen Untersuchungsmerkmale der Schülerinnen und Schüler fallen die Mittelwerte der Gruppe Abbrecher signifikant geringer aus als die Einschätzungen der Gruppe Teilnehmer. Mit Blick auf die Effektstärke kann ein großer Effekt jedoch nur beim Merkmal der Leistungsunterscheidung festgestellt werden. Demnach nehmen Schüler und Schülerinnen, die im Mittel eine höhere Mathematikleistung erzielen, weiterhin an der Studie teil.

6.1.2 Drop-Out-Analyse der Studie 2

Die Studie 2 umfasste zum ersten Messzeitpunkt Daten von 171 Lehrkräften. Von diesen nahmen 87 Lehrkräfte nur einmalig an der Befragung teil (Abbrecher), 51 Personen füllten die Materialien erst zum zweiten Messzeitpunkt aus (hier nicht beachtet) und von 84 Lehrkräften liegen Daten zu beiden Messzeitpunkten vor (Teilnehmer).

Tabelle 27 gibt einen Überblick über die Charakteristika der Personen in den Gruppen Abbrecher und Teilnehmer. Für die in der Studie erhobenen Merkmale ergaben sich keine bedeutenden Unterschiede zwischen den Lehrkräften, die die Untersuchung fortführten und denen, die abbrachen. Gründe für den Abbruch wurden in der Projektlaufzeit nicht bekannt.

Tabelle 27. Deskriptive Statistiken (in Prozent) für die Lehrergruppen Teilnehmer und Abbrecher in Studie 2

Merkmal		Abbrecher (n=87)	Teilnehmer (n=84)	χ^2	df	p
Gruppe	Projektschule	45.5	54.5	1.264	1	.261
	Vergleichsschulen	54.3	45.7			
Bundesland	Berlin/BB	46.8	53.2	0.215	1	.643
	Hessen	50.5	49.5			
Geschlecht	männlich	17.3	82.7	31.659	1	.000
	weiblich	64.3	35.7			
Alter	unter 25	100.0	0.0	3.498	5	.624
	25-29	42.9	57.1			
	30-39	48.5	51.5			
	40-49	43.4	56.6			
	50-59	52.3	47.7			
	60 und älter	71.4	28.6			
Schulform	Gymnasium	46.1	53.9	0.693	1	.405
	Andere	52.4	47.6			

Systematische Unterschiede sind zwischen den beiden Gruppen der Abbrecher und Teilnehmer mit Ausnahme des Merkmals Geschlecht nicht auszumachen. So nehmen mehr als die Hälfte der Lehrkräfte in den Projektschulen an weiteren Testungen teil, in den Vergleichsschulen steigen mehr als die Hälfte der Lehrkräfte nach der ersten Testung aus. Weiterhin führen von den Lehrkräften des zusammengefassten Bundeslandes Berlin/Brandenburg mehr als die Hälfte das Projekt fort, in Hessen brechen mehr als die Hälfte ab. Die altersbedingte Teilnahme konzentriert sich auf die Lehrkräfte, die bereits seit einigen Jahren im Schuldienst tätig sind oder noch mehr als 5 Jahre im Schuldienst tätig sein werden (25-59 Jahre). Von den Gymnasiallehrkräften entscheiden sich mehr als

die Hälfte, das Projekt weiterzuverfolgen. In den anderen Schulformen fällt die Entscheidung für mehr als die Hälfte negativ aus. Statistisch bilden sich die Gruppen der Teilnehmer und der Abbrecher jedoch zufällig.

In der Schülerstichprobe der Studie 2 zeigen sich ähnliche Ergebnisse. Von den 2.458 Schülerinnen und Schüler der ersten Erhebung nahmen 1.970 auch an der zweiten Erhebung teil. Damit verbleiben 488 Lernerinnen und Lernern, von denen keine Leistungs- und Befragungsdaten zum zweiten Messzeitpunkt erhoben werden konnten (Abbrecher). Eine Charakterisierung der 488 Abbrecher ist Tabelle 28 zu entnehmen. Es zeigt sich, dass bei einer Betrachtung der Zugehörigkeit zu Projekt- versus Vergleichsschulen keine bedeutenden Unterschiede zu beobachten sind, d.h. eine Nicht-Teilnahme an den Testungen und Befragungen kann nicht auf eine Teilnahme am Projekt zurückgeführt werden.

Tabelle 28. Mittelwertunterschiede zwischen Teilnehmern und Abbrechern in Studie 2, Schülerstichprobe

Merkmal	Gruppe	<i>MW (SD)</i>	<i>F</i>	<i>p</i>	η^2
Selbstkonzept	Teilnehmer	2.49 (0.78)	6.76	.009	.003
	Abbrecher	2.38 (0.79)			
Interesse	Teilnehmer	2.10 (0.72)	0.75	.387	.000
	Abbrecher	2.07 (0.76)			
Lehrerwahrnehmung	Teilnehmer	2.89 (0.80)	5.26	.022	.002
	Abbrecher	2.79 (0.80)			
Argumentieren	Teilnehmer	3.87 (1.02)	0.51	.475	.000
	Abbrecher	3.83 (1.08)			
Probleme Lösen	Teilnehmer	3.77 (0.95)	0.00	.970	.000
	Abbrecher	3.77 (0.99)			
Modellieren	Teilnehmer	3.40 (1.08)	0.88	.349	.000
	Abbrecher	3.45 (1.09)			
Darstellungen verwenden	Teilnehmer	3.39 (0.91)	0.46	.496	.000
	Abbrecher	3.36 (0.97)			
technisch Arbeiten	Teilnehmer	3.89 (0.82)	4.33	.038	.002
	Abbrecher	3.80 (0.89)			
Kommunizieren	Teilnehmer	3.89 (1.00)	0.09	.763	.000
	Abbrecher	3.87 (1.06)			
Mathematikleistung	Teilnehmer	526.11 (103.67)	19.28	< .001	.008
	Abbrecher	503.40 (96.62)			

6.2 Forschungsfrage 1: Merkmalsveränderungen und Entwicklungstendenzen von Mathematiklehrkräften

Formal untersucht die Forschungsfrage 1 die Veränderung von Untersuchungsmerkmalen der Probanden während der Anwendung einer Intervention. Bezogen auf die Einführung der Bildungsstandards bedeutet die Analyse der Untersuchungsmerkmale, inwiefern eine

unterstützende Begleitung des Unterrichts eine bedeutende Entwicklung in der Unterrichtsdurchführung im Sinne einer Kompetenzorientierung hervorruft. Es soll unter Beachtung des Ausgangsniveaus untersucht werden, wie die Lehrkräfte in den relevanten Merkmalen ihre Einstellungen und Unterrichtsaktivitäten im Vergleich zu einer Gruppe ohne begleitende Unterstützung verändern. Dazu werden die Untersuchungsmerkmale zunächst im Verlauf dargestellt und auf Gruppenunterschiede untersucht. Die längsschnittliche Anlage der Studie 1 macht eine Gegenüberstellung der Daten im Querschnitt, Längsschnitt und als imputierter Datensatz sinnvoll. Die Merkmale werden dabei einzeln betrachtet. Abschließend erfolgt jeweils eine Modellierung der Daten unter Beachtung der zeitlichen Abstände zwischen den Erhebungen mit dem Fokus auf Wachstumseffekte. Diese Mehrebenenmodellierung wird nur unter Verwendung des imputierten Datensatzes durchgeführt.

6.2.1 Studie 1 „Erprobung eines Konzeptes zur Implementation der Bildungsstandards in Berliner und Brandenburger Schulen“

Die Darstellung der Merkmalsentwicklungen auf Grundlage der Rohdaten erfolgt in der Gegenüberstellung der gesamten Stichprobe und der Lehrkräfte, die an allen drei Messzeitpunkten die Untersuchungsmaterialien bearbeiteten. Dieses Vorgehen wurde gewählt, da teilweise auffallende Unterschiede in den Verläufen deutlich werden. Weiterhin kann vor dem Hintergrund der Gegenüberstellung abgewogen werden, ob es sich bei den Probanden des Längsschnitt um eine Negativauswahl handelt. Für die folgenden deskriptiven Analysen gelten die Teilnehmerzahlen aus Tabelle 29.

Tabelle 29. Teilnehmeranzahl der Lehrkräfte über die Messzeitpunkte

	Querschnitt		Längsschnitt		imputierter Datensatz	
	PS	VS	PS	VS	PS	VS
1. MZP	51 - 52*	59 - 60*	16	28	26	46
2. MZP	25 - 27*	52 - 55*	16	28	26	46
3. MZP	22 - 25*	47 - 48*	16	28	26	46

Anmerkung. PS = Projektschulen, VS = Vergleichsschulen, * Teilnehmeranzahl variiert je nach Merkmal

Durch die Imputation nach festgelegten Kriterien (vgl. Abschnitt 5.4.5) konnten Datensätze von 26 Lehrkräften in den Projektschulen und 46 Lehrkräften in den Vergleichsschulen generiert werden.

6.2.1.1 Einstellungen gegenüber der Einführung von nationalen Bildungsstandards

Hinsichtlich der positiven Einstellungen gegenüber den Bildungsstandards stellen sich die Lehrkräfte der Projekt- und Vergleichsschulen als gleichwertig dar. Zwischen den Grup-

pen sind zum ersten Erhebungszeitpunkt keine Unterschiede zu beobachten. Die Skalenergebnisse liegen durchgängig im positiven Bereich. Eine signifikante Veränderung ist nur für die Lehrkräfte des Längsschnitts zu beobachten. Hier entwickeln die Lehrkräfte der Projektschulen ihre positiven Einstellungen signifikant stärker als die Lehrkräfte in den Vergleichsschulen, was auf die stetige Arbeit im Projekt zurückgeführt wird.

Tabelle 30. Deskriptive Statistiken zum Merkmal positive Einstellungen gegenüber den Bildungsstandards

Merkmal	MZP 1				MZP 2				MZP 3			
	<i>N</i>	<i>MW</i> (<i>SD</i>)	<i>F</i>	η^2	<i>N</i>	<i>MW</i> (<i>SD</i>)	<i>F</i>	η^2	<i>N</i>	<i>MW</i> (<i>SD</i>)	<i>F</i>	η^2
positive Einstellungen (Querschnitt)												
Projektschulen	52	3.46 (0.70)			25	3.55 (0.74)			22	3.52 (0.84)		
Vergleichsschulen	60	3.33 (0.78)	0.84	.01	52	3.57 (0.72)	0.02	.00	47	3.46 (0.54)	0.11	.00
gesamt	112	3.39 (0.74)			77	3.56 (0.72)			69	3.48 (0.64)		
positive Einstellungen (Längsschnitt)												
Projektschulen	16	3.52 (0.77)			15	3.72 (0.58)			14	3.83 (0.40)		
Vergleichsschulen	28	3.51 (0.81)	0.00	.00	28	3.57 (0.74)	0.43	.01	28	3.47 (0.57)	4.38*	.10
gesamt	44	3.51 (0.79)			43	3.62 (0.69)			42	3.59 (0.54)		
positive Einstellungen (imputiert)												
Projektschulen	26	3.58 (0.72)			26	3.54 (0.73)			26	3.67 (0.44)		
Vergleichsschulen	46	3.46 (0.73)	0.49	.01	46	3.55 (0.70)	0.02	.00	46	3.49 (0.51)	2.28	.03
gesamt	72	3.50 (0.73)			72	3.55 (0.71)			72	3.56 (0.48)		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen

Die Gegenüberstellung der Daten des Querschnitts mit denen des Längsschnitts zeigt grafisch große Differenzen. Alle Lehrkräfte des Längsschnitts zeigen von Anfang an eine etwas höher ausgeprägte positive Einstellung gegenüber den Bildungsstandards. Die Lehrkräfte in den Projektschulen schaffen es über die Projektlaufzeit, ihre positiven Einstellungen weiter zu entwickeln. Die Schere zum 3. Messzeitpunkt zwischen den Projekt- und Vergleichsschulen geht eindeutig stärker auseinander als in den Daten des Querschnitts. Folglich beteiligten sich zu den einzelnen Messungen der Projektschulen Lehrkräfte, deren Einstellungen weniger positiv ausgeprägt waren und deren Daten den Mittelwert für die Projektschulen nach unten absinken ließ (Querschnitt). Anders formuliert verblieben diejenigen Lehrkräfte im Projekt, deren Einstellung gegenüber den Bildungsstandards positiv war und die durch die Arbeit an einem kompetenzorientierten Unterricht ihre positive Einstellung noch weiter ausbauen konnten (vgl. Abbildung 16).

Statistisch kann durch eine einfaktorielle Varianzanalyse (ANOVA) nachgewiesen werden, dass sich die Gruppen (Projektgruppe und Kontrollgruppe) in ihren Ausgangswerten sowohl im Querschnitt als auch im Längsschnitt und dem imputierten Datensatz nicht signifikant unterscheiden.

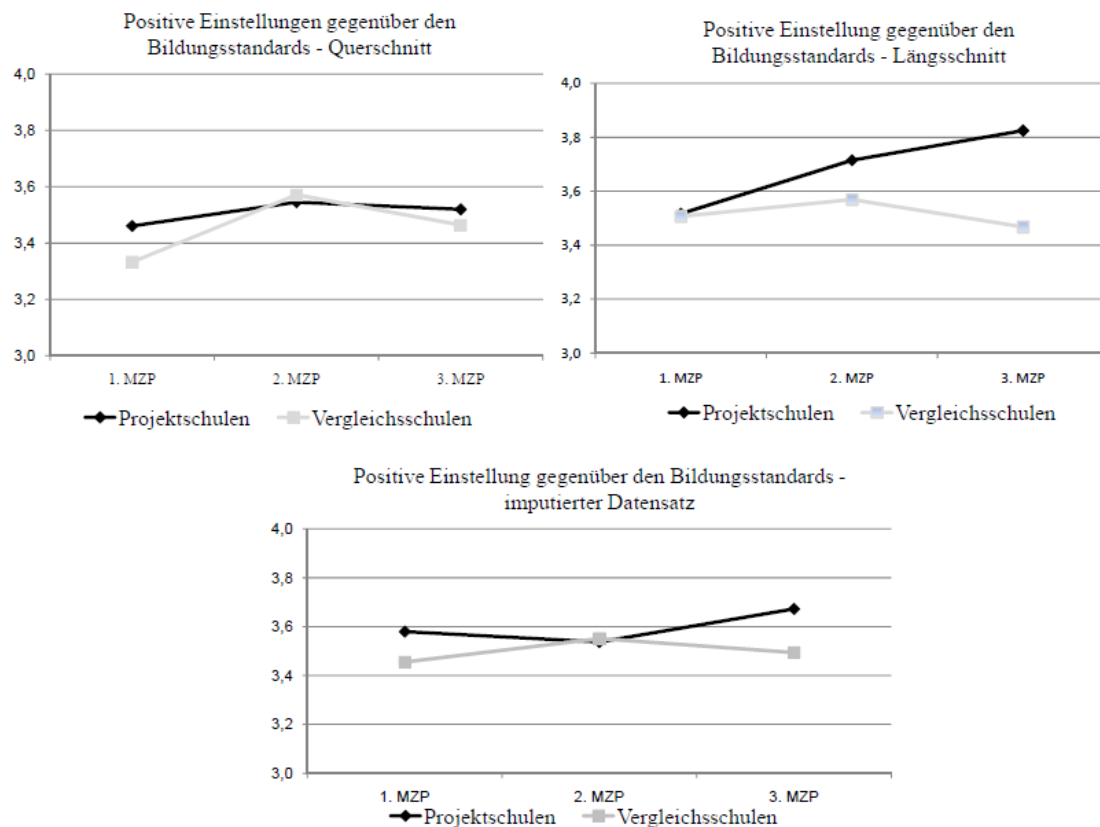


Abbildung 16. Positive Einstellungen gegenüber den Bildungsstandards im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

Zur Analyse von systematischen Gruppeneffekten über die Projektlaufzeit wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor gerechnet. Es konnten jedoch keine signifikanten Haupt- und Interaktionseffekte nachgewiesen, womit die Entwicklung der positiven Einstellungen gegenüber den Bildungsstandards zufällig bleibt ($F(2,80)=0.30$; $p > .05$).

Im intraindividuellen Vergleich kann eine größere Streuung der Werte in den Ausgangsdaten der Vergleichsschulen beobachtet werden, die sich zum 3. Messzeitpunkt verringert. Tendenziell ist für die Mehrheit der Lehrkräfte in den Vergleichsschulen ein leichter Abfall der positiven Einstellungen zu erkennen, wenngleich dem einige positive Entwicklungen gegenüberstehen.

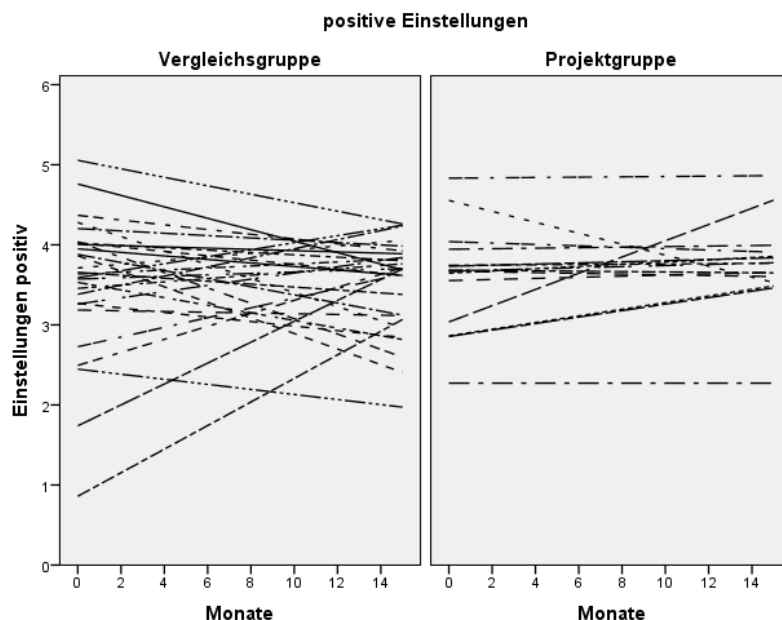


Abbildung 17. Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der positiven Einstellungen gegenüber den Bildungsstandards.

Eine Regression der Zeitvariable (mit unterschiedlichen Zeitabständen) auf die abhängige Variable der positiven Einstellungen für jede Lehrkraft zeigt in den Projektschulen keine signifikanten Ergebnisse. Lediglich in den Vergleichsschulen bauen zwei Lehrkräfte ihre positiven Einstellungen schwach signifikant ab und zwei Lehrkräfte schwach signifikant auf ($p < .10$).

Der durchschnittliche Ausgangswert für Lehrkräfte der Längsschnittstichprobe liegt bei $M=3.59^{55}$. Es ist ein durchschnittliches Wachstum pro Monat von $b = .002$ Skalenpunkten zu beobachten. Eine starke negative Korrelation zwischen dem Ausgangswert und der Steigung deutet darauf hin, dass Lehrkräfte mit höheren positiven Einstellungen zu Beginn der Studie sich signifikant langsamer entwickeln als Lehrkräfte mit geringer ausgeprägten positiven Einstellungen (Tabelle 31).

Tabelle 31. Deskriptive Statistik für den individuellen Wachstumsparameter aus individuellen Regressionen für positive Einstellungen gegenüber den Bildungsstandards als eine Funktion der Zeit ($n=42$)

	Konstante (Ausgangswert)	Steigung (Veränderungsrate)
Mittelwert	3.59	0.00
Standardabweichung	0.78	0.05
Korrelation		-.76**

⁵⁵ Der Wert unterscheidet sich vom Mittelwert in Tabelle 34 für die Stichprobe im Längsschnitt, da bei der Schätzung der Konstanten die unterschiedlichen Zeitabstände zwischen den Befragungen in Rechnung gestellt wurden (siehe Zeitvariable). Es ist hier allerdings nicht möglich, eine Gruppenunterscheidung zur differenzierten Analyse durchzuführen. Dazu sei auf die Mehrebenenmodelle in Tabelle 34 verwiesen.

Hinsichtlich der negativen Einstellungen gegenüber der Einführung von Bildungsstandards zeigt sich zunächst, dass sich die Gruppen in ihrer Ausgangssituation nicht voneinander unterscheiden. Im Verlauf des Projekts können die Projektschulen ihre negativen Einstellungen abbauen, was sich insbesondere in den Querschnittsdaten widerspiegelt (Tabelle 32). Die Kolleginnen und Kollegen der Vergleichsschulen verharren nahezu auf dem Anfangsniveau.

Tabelle 32. Deskriptive Statistiken zu dem Merkmal negative Einstellungen gegenüber den Bildungsstandards

Merkmal	MZP 1				MZP 2				MZP 3			
	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2
negative Einstellungen (Querschnitt)												
Projektgruppe	52	2.61 (0.65)			25	2.39 (0.56)			22	2.18 (0.44)		
Vergleichsgruppe	60	2.69 (0.63)	0.47	.00	52	2.69 (0.64)	4.06*	.05	47	2.67 (0.53)	14.10***	.17
gesamt	112	2.65 (0.64)			77	2.59 (0.62)			69	2.51 (0.55)		
negative Einstellungen (Längsschnitt)												
Projektgruppe	16	2.29 (0.42)			15	2.19 (0.47)			14	2.09 (0.41)		
Vergleichsgruppe	28	2.55 (0.65)	1.98	.05	28	2.53 (0.62)	3.42~	.08	28	2.55 (0.57)	7.12*	.15
gesamt	44	2.46 (0.59)			43	2.41 (0.59)			42	2.39 (0.56)		
negative Einstellungen (imputiert)												
Projektgruppe	26	2.33 (0.43)			26	2.34 (0.50)			26	2.22 (0.36)		
Vergleichsgruppe	46	2.60 (0.59)	4.22	.06	46	2.57 (0.56)	3.17	.04	46	2.56 (0.49)	9.84	.12
gesamt	72	2.50 (0.55)			72	2.49 (0.55)			72	2.44 (0.48)		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Grafisch zeigen die Lehrkräfte des Längsschnitts sowohl in den Vergleichs- als auch in den Projektschulen eine deutlich geringer ausgeprägte negative Einstellung gegenüber den Bildungsstandards, im Vergleich zur Betrachtung im Querschnitt. In den Vergleichsschulen ändert sich jedoch über die Laufzeit nichts an deren Einstellungen, die Lehrkräfte in den Projektschulen bauen jedoch ihre negativen Einstellungen weiter ab (Abbildung 18). Folglich nahmen Lehrkräfte in beiden Gruppen nur einmalig an der Studie teil, die den Bildungsstandards deutlich kritischer gegenüberstehen und eine weitere Teilnahme verweigerten. Die grafisch auffallenden Entwicklungen ergaben sich bei einer statistischen Überprüfung jedoch als zufällig, d.h. der grafisch erkennbare starke Abbau der negativen Einstellungen gegenüber den Bildungsstandards in den Projektschulen kann statistisch nicht nachgewiesen werden ($F(2,80)=0.17$; $p > .05$).

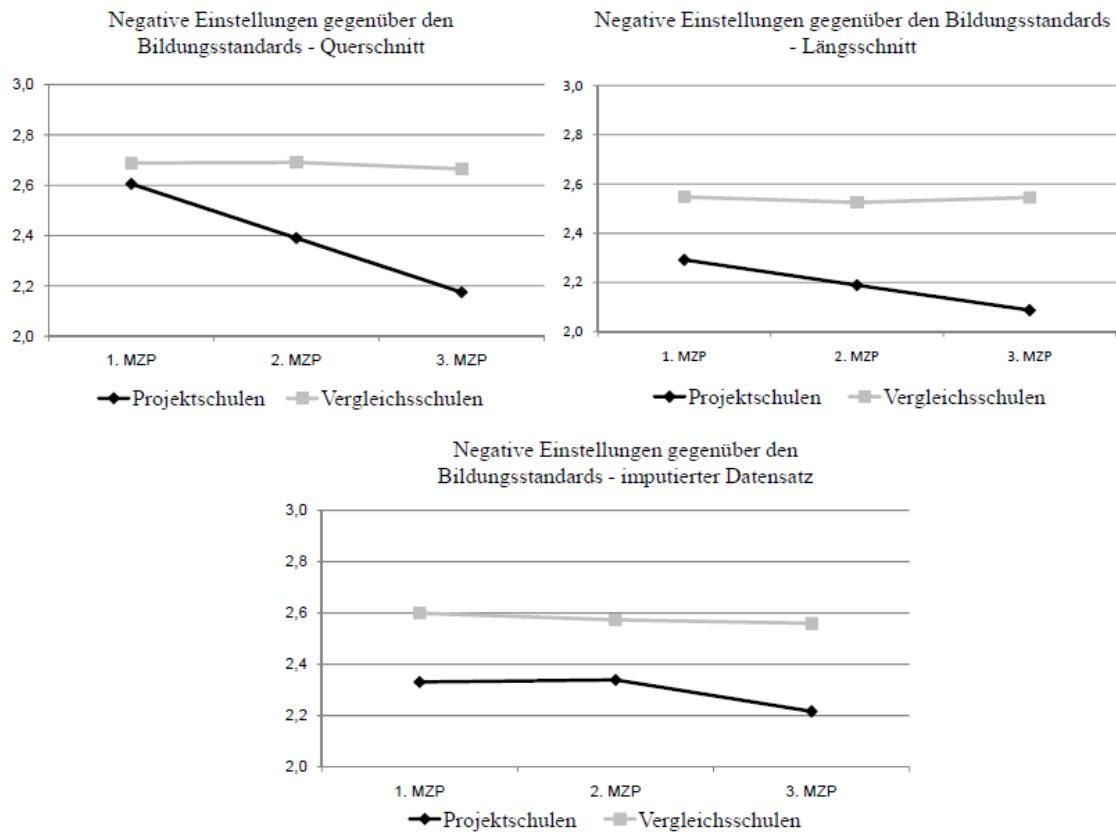


Abbildung 18. Negative Einstellungen gegenüber den Bildungsstandards im Verlauf; Gegenüberstellung gesamte Stichprobe und Längsschnitt.

In der intraindividuellen Betrachtung können mit einer Ausnahme ebenfalls keine signifikanten Entwicklungen ausgemacht werden. Einzig eine Lehrkraft aus den Vergleichsschulen kann ihre negativen Einstellungen signifikant im Verlauf des Projekts abbauen. Die anderen Entwicklungen unterschreiten die Zufallswahrscheinlichkeit von $p < .05$ nicht.

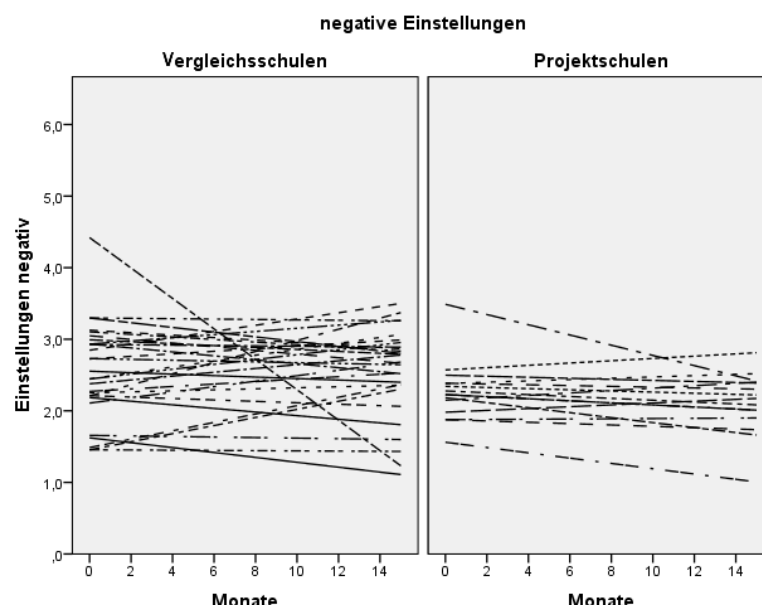


Abbildung 19. Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der negativen Einstellungen gegenüber den Bildungsstandards.

Im Durchschnitt kann für dieses Merkmal unter Berücksichtigung der unterschiedlichen Testzeitpunkte ein Anfangswert von $M=2.45$ beobachtet werden. Die Lehrkräfte bauen im Durchschnitt ihre negativen Einstellungen pro Monat in Höhe von $b=-0.004$ ab. Der signifikant negative Zusammenhang mittlerer Stärke zwischen Ausgangsniveau und Steigung weist darauf hin, dass sehr skeptische Lehrkräfte mit hohen Werten in der Konstante, ihre Befürchtungen auch langsamer abbauen als Lehrkräfte, deren negative Einstellungen zu Beginn weniger hoch ausgeprägt waren.

Tabelle 33. Deskriptive Statistik für den individuellen Wachstumsparameter aus individuellen Regressionen für negative Einstellungen gegenüber den Bildungsstandards als eine Funktion der Zeit ($n=43$)

	Konstante (Ausgangswert)	Steigung (Veränderungsrate)
Mittelwert	2.45	0.00
Standardabweichung	0.61	0.04
Korrelation		-.55**

Anmerkung. Exakter Wert der Steigung (Veränderungsrate) = -0.0038. Die Veränderungsrate in diesem Merkmal ist damit negativ.

Die Berechnung eines durchschnittlichen Anfangswertes und einer Steigerungsrate anhand individueller linearer Regressionen wurde für jedes Merkmal durchgeführt. Alle Merkmale weisen eine signifikant negative Korrelation auf, womit für jedes Merkmal gilt: Ein höherer durchschnittlicher Skalenwert im Ausgangswert führt zu einer langsame-

ren Steigung (Entwicklung)⁵⁶. Diese Aussage ergibt sich für alle untersuchten Merkmale und wird daher in den weiteren Ausführungen vernachlässigt⁵⁷.

Die finalen Mehrebenenmodelle führen zur Beantwortung der Forschungsfragen. Hinsichtlich der Einstellungen der Lehrkräfte gegenüber den Bildungsstandards ergibt sich folgendes Bild. In der Ausgangssituation unterscheiden sich die Lehrkräfte der Projektschulen nur hinsichtlich ihrer negativen Einstellung von den Lehrkräften in den Vergleichsschulen. Als hinreichende Teilnahmebedingung zur Arbeit im Projekt stehen die Lehrkräfte der Projektschulen den Bildungsstandards um 0.247 Skalenpunkte weniger kritisch gegenüber als die Lehrkräfte in den Vergleichsschulen.

Unter Berücksichtigung der unterschiedlichen Zeitabstände zwischen den Messungen können jedoch keine signifikanten Entwicklungen insbesondere bei den Lehrkräften der Projektschulen nachgewiesen werden. Es ist zu beobachten, dass die Lehrkräfte in den Vergleichsschulen ihre positiven Einstellungen monatlich um .005 Skalenpunkte entwickeln und die negativen Einstellungen um .003 Skalenpunkte abbauen. Die Lehrkräfte in den Projektschulen bauen ihre negativen Einstellungen in einem etwas höheren Maßstab ab (-.005 Skalenpunkte), die Unterschiede weisen jedoch auch mit Blick auf die hohen Standardfehler keine Signifikanz auf.

Tabelle 34. Wachstumseffekte (Modell C) der Einstellungen gegenüber der Einführung von nationalen Bildungsstandards, Studie 1, Lehrkräfte

		Parameter	positive Einstellungen	negative Einstellungen
<i>Feste Effekte</i>				
Ausgangswert	Konstante	γ_{00}	3.466*** (.118)	2.601*** (.093)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	.080 (.183)	-.247~ (.129)
Wachstumswert	Konstante	γ_{10}	.005 (.009)	-.003 (.007)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	.002 (.014)	-.005 (.010)
<i>Zufallseffekte</i>				
Level 1	innerhalb Personen	σ^2_{ϵ}	.201	.111
Level 2	Ausgangswert	σ^2_0	.393***	.204***
	Wachstumswert	σ^2_1	.001*	.001*

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen;

HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

⁵⁶ Z.B. Lehrkräfte, die zum Projektbeginn in ihren Klassen häufiger mathematisch Argumentieren, steigern sich in der Ausprägung dieser Kompetenz bis zum Projektende langsamer.

⁵⁷ Die Korrelationstabellen für jedes Merkmal sind in Anhang C einsehbar.

Die Analyse der Varianzanteile auf den verschiedenen Ebenen des Wachstumsmodells weist darauf hin, dass auf der intraindividuellen Ebene keine signifikante Restvarianz verbleibt, die es zu erklären gilt. Mit Blick auf die so genannten Nullmodelle (siehe Anhang D) konnte die Varianz jedoch reduziert werden⁵⁸.

6.2.1.2 Zusammenarbeit und Kooperation in der Fachkonferenz

Die Zusammenarbeit in der Fachkonferenz Mathematik unterliegt in der Praxis schulinternen Bedingungen und variiert je nach Stand der Schulentwicklungsmaßnahmen von Schule zu Schule. So spielen räumliche (evt. mehrere Schulstandorte) und organisatorische Aspekte (Stundenpläne der Lehrkräfte, Anzahl der unterrichteten Stunden im Fach, Belastung durch andere Schulentwicklungsmaßnahmen, Teilzeitarbeit) eine wichtige Rolle in der Ausprägung der Zusammenarbeit. Das zeigt sich auch in den unterschiedlichen Anfangsniveaus der Kontroll- und Projektschulen. Sie unterscheiden sich signifikant voneinander, wobei die Vergleichsschulen ein höheres Maß an Zusammenarbeit aufweisen als die Projektschulen. Tabelle 35 stellt die verschiedenen Entwicklungen gegenüber.

Tabelle 35. Deskriptive Statistiken zum Merkmal Zusammenarbeit in der Fachkonferenz

Merkmal	MZP 1				MZP 2				MZP 3			
	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>F</i>	η^2
Fachkonferenz (Querschnitt)												
Projektgruppe	51	2.23 (0.40)			27	2.48 (0.50)			25	2.51 (0.43)		
Vergleichsgruppe	60	2.44 (0.36)	8.48**	.07	55	2.59 (0.57)	0.73	.01	47	2.66 (0.56)	1.39	.02
gesamt	111	2.34 (0.40)			82	2.55 (0.55)			72	2.61 (0.52)		
Fachkonferenz (Längsschnitt)												
Projektgruppe	16	2.20 (0.47)			16	2.41 (0.50)			16	2.44 (0.42)		
Vergleichsgruppe	28	2.55 (0.34)	8.18**	.16	28	2.53 (0.49)	0.58	.01	28	2.60 (0.37)	1.87	.04
gesamt	44	2.42 (0.42)			44	2.49 (0.49)			44	2.54 (0.39)		
Fachkonferenz (imputiert)												
Projektgruppe	26	2.25 (0.47)			26	2.44 (0.47)			26	2.56 (0.40)		
Vergleichsgruppe	46	2.49 (0.38)	5.84~	.08	46	2.63 (0.54)	2.23	.03	46	2.60 (0.46)	0.32	.01
gesamt	72	2.40 (0.43)			72	2.56 (0.53)			72	2.58 (0.44)		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD); F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen

Mit einem signifikant kleineren Mittelwert von $MW = 2.23$ stellen die Projektschulen in Bezug auf dieses Merkmal eine Negativauswahl dar ($F(1,42) = 8.48$; $p = .004$, $\eta^2 = .07$).

⁵⁸ Bereits in den Nullmodellen konnte keine zu erklärende Varianz auf Ebene eins berechnet werden.

Im Querschnitt entwickeln sich die Lehrkräfte der Projektschulen stärker als die Lehrkräfte in den Vergleichsschulen, da die zu Beginn des Projektes noch vorliegenden signifikanten Unterschiede ausgeglichen werden. Diese Entwicklung kann darin begründet sein, dass Lehrkräfte der Projektschulen mit einer geringen Bewertung der Zusammenarbeit in der Fachkonferenz aus dem Projekt ausgestiegen sind. Diese Aussage wird validiert durch die Betrachtung der Längsschnittdaten. Hier zeigt sich, dass die Lehrkräfte, die durchgängig am Projekt beteiligt waren, eine Entwicklung von 0.21 Skalenpunkten zum zweiten Messzeitpunkt aufweisen. Folglich sind im Querschnitt all diejenigen aus dem Projekt ausgestiegen, die eine Zusammenarbeit eher schlechter bewerten. Auch zum dritten Messzeitpunkt konnten die Projektschulen ihre Zusammenarbeit noch einmal stärken, wenngleich der Mittelwert am Ende des Projekts noch immer unterhalb der Vergleichsschulen lag.

Auch in den Vergleichsschulen brachen die Lehrkräfte das Projekt ab, deren Zusammenarbeit in der Fachkonferenz weniger stark ausgeprägt war. Dieser Rückschluss ergibt sich bei einem grafischen Vergleich zum Längsschnitt. Hier zeigt sich, dass die Ausprägung der Lehrkräfte in den Vergleichsschulen, die durchgängig am Projekt teilnahmen (Längsschnitt), höher liegt als die Ausprägung zum 1. Messzeitpunkt bei der Betrachtung im Querschnitt.

Eine einfaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor ergab einen Haupteffekt für die Zeit ($F(2, 84) = 3.71; p < .05$) und einen schwachen Haupteffekt für die Gruppe ($F(1, 42) = 3.26; p < .10$). Das Merkmal der Zusammenarbeit zeigt demnach signifikante Entwicklungen; die Gruppen entwickeln sich insgesamt über die Zeit und auch die Gruppen unterscheiden sich voneinander. Jedoch konnte ein Interaktionseffekt, der die unterschiedliche Entwicklung der Gruppen über die Zeit abbildet, nicht nachgewiesen werden ($F(2, 84) = 2.36; p > .05$).

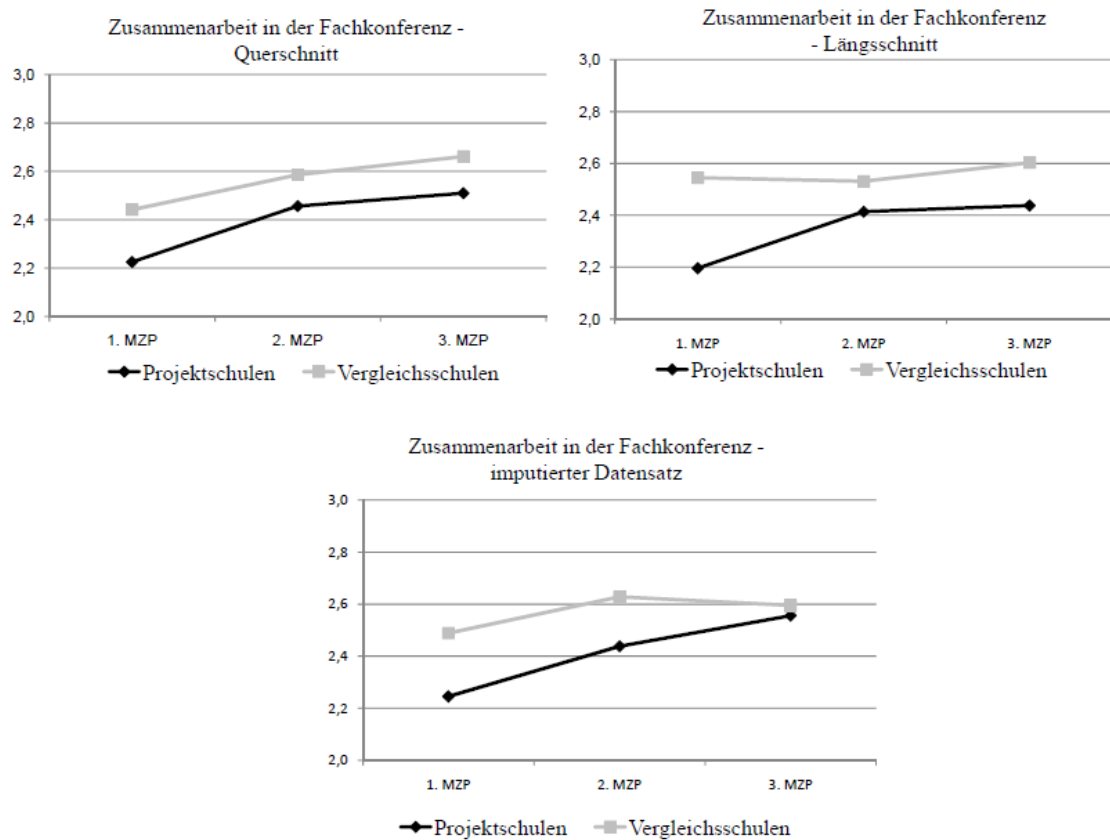


Abbildung 20. Zusammenarbeit in der Fachkonferenz im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

Die intraindividuellen Entwicklungen bilden ein vergleichsweise homogenes Bild. In den Vergleichsschulen erhöht sich im Verlauf des Projektes die Streuung in der Wahrnehmung der Zusammenarbeit, in den Projektschulen zeigt sich eine gegensätzliche Entwicklung. Zusätzlich ist in den Projektschulen tendenziell eine positive Entwicklung zu erkennen, während die Mehrheit der Lehrkräfte in den Vergleichsschulen auf dem anfänglichen Niveau verharrt.

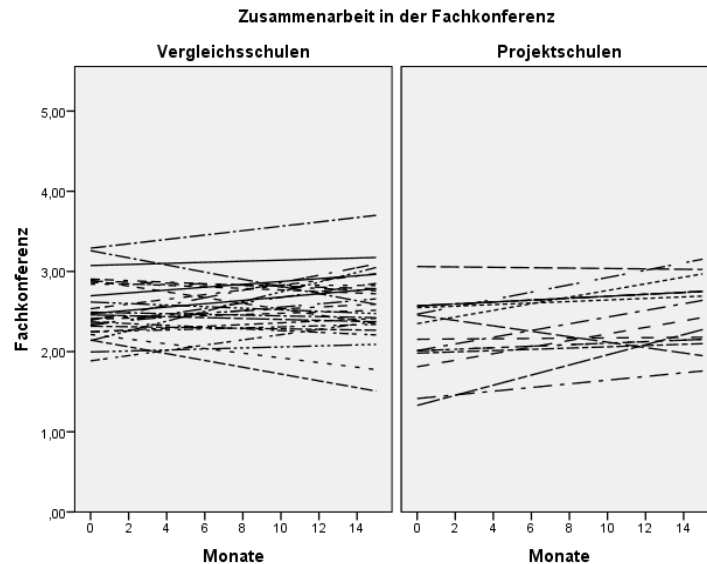


Abbildung 21. Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der Zusammenarbeit in der Fachkonferenz.

Das Merkmal Kooperation in der Fachkonferenz bezieht sich auf die Organisation und das Klima in der Fachkonferenz. Auch dieses Merkmal ist von schulinternen Bedingungen abhängig, wobei die Tiefe der Zusammenarbeit hier keine Rolle spielt. Tabelle 36 zeigt hier ebenfalls, dass die Vergleichsschulen in ihren Ausgangswerten auf einem höheren Niveau miteinander kooperieren als die Projektschulen. Die Unterschiede in den Ausgangswerten sind nicht signifikant ($F(1,42) = 1.56, p > .05$) und damit zufällig. Im Verlauf der Studie zeigt sich jedoch eine sprunghafte Steigerung der Kooperation in den Projektschulen, während die Lehrkräfte in den Vergleichsschulen auf ihrem Niveau nahezu verharren bzw. leicht absinken. Die Projektschulen steigerten ihre Kooperation und verbleiben am Ende der Studie auf einem höheren Niveau.

Tabelle 36. Deskriptive Statistik zum Merkmal Kooperation in der Fachkonferenz

Merkmal	MZIP 1				MZIP 2				MZIP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
Kooperation (Querschnitt)												
Projektgruppe	51	2.84 (0.67)			27	3.18 (0.60)			25	3.19 (0.60)		
Vergleichsgruppe	60	3.03 (0.49)	2.72	.02	54	3.06 (0.45)	0.88	.01	48	3.00 (0.50)	1.87	.03
gesamt	111	2.94 (0.58)			81	3.10 (0.51)			73	3.07 (0.54)		
Kooperation (Längsschnitt)												
Projektgruppe	16	2.98 (0.75)			16	3.29 (0.63)			16	3.26 (0.61)		
Vergleichsgruppe	28	3.19 (0.41)	1.56	.04	28	3.17 (0.42)	0.49	.01	28	3.07 (0.54)	1.20	.03
gesamt	44	3.12 (0.56)			44	3.21 (0.51)			44	3.14 (0.57)		
Kooperation (imputiert)												
Projektgruppe	26	3.03 (0.65)			26	3.26 (0.55)			26	3.19 (0.58)		
Vergleichsgruppe	46	3.04 (0.47)	0.05	.00	46	3.08 (0.45)	2.18	.03	46	2.95 (0.51)	3.54	.05
gesamt	72	3.04 (0.54)			72	3.14 (0.49)			72	3.04 (0.55)		

Anmerkungen: $\sim p < .10$; $*p < .05$; $**p < .01$; $***p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen

Die von den Lehrkräften der Projektschulen wahrgenommene Steigerung in der Kooperation kann auf die mit der Projektarbeit verbundenen Treffen der Fachkonferenz zurückgeführt werden.

Diese Entwicklung der Projektschulen kann auch statistisch belegt werden. Es ergibt sich ein signifikanter Haupteffekt für die Zeit ($F(2,82) = 3.45$; $p = .036$) und ein noch bedeutenderer Interaktionseffekt ($F(2,82) = 7.80$; $p = .001$). Damit kann gezeigt werden, dass sich die Projektschulen in ihrer Kooperation in der Fachkonferenz systematisch stärker entwickeln als die Vergleichsschulen.

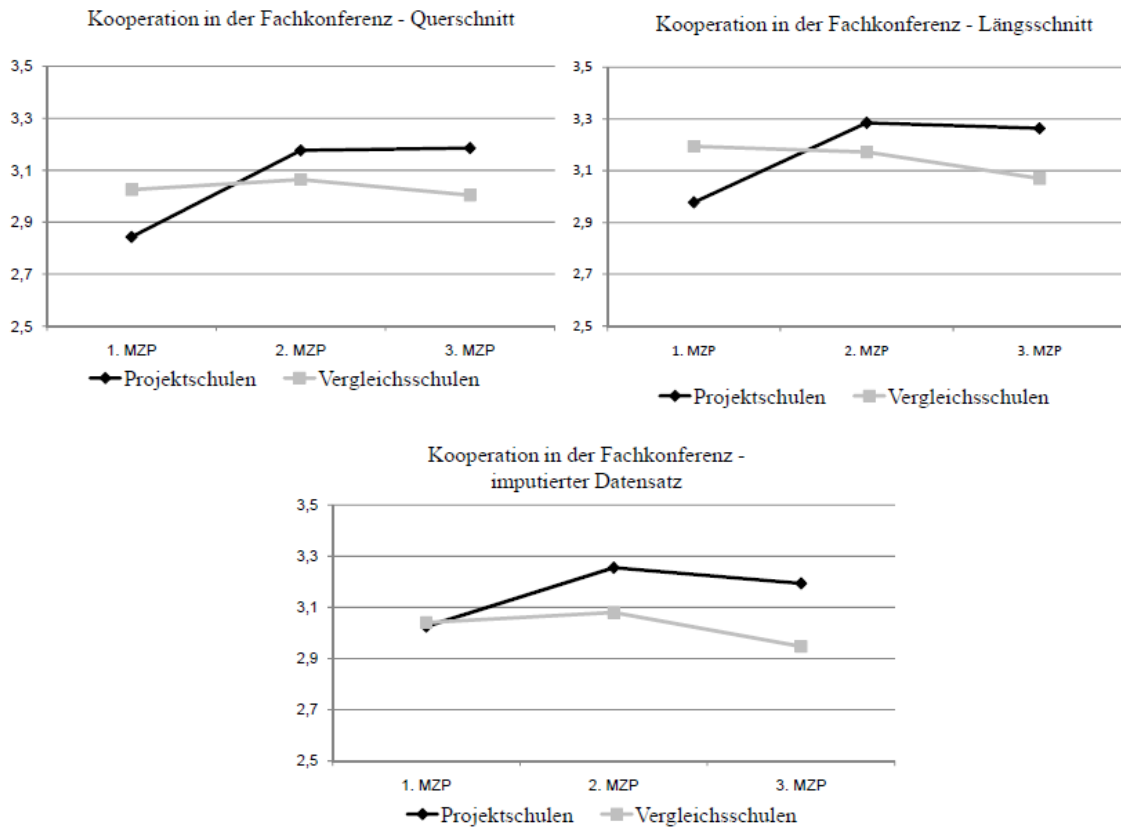


Abbildung 22. Kooperation in der Fachkonferenz im Verlauf; Gegenüberstellung Querschnitt, Längsschnitt und imputierte Daten.

Auch in der intraindividuellen Entwicklung ist diese positive Entwicklung erkennbar. Mehrere Lehrkräfte in den Projektschulen weisen starke Steigerungen im Verlauf des Projektes auf. Mit Ausnahme von drei Lehrkräften ist ein positiver Entwicklungstrend in den Projektschulen zu beobachten. Die Mehrheit der Lehrkräfte in den Vergleichsschulen verharrt auf ihrem Anfangsniveau. Auffällig sind auch die mindestens drei starken negativen Entwicklungen von Lehrkräften in den Vergleichsschulen, deren Gründe jedoch spekulativ wären.

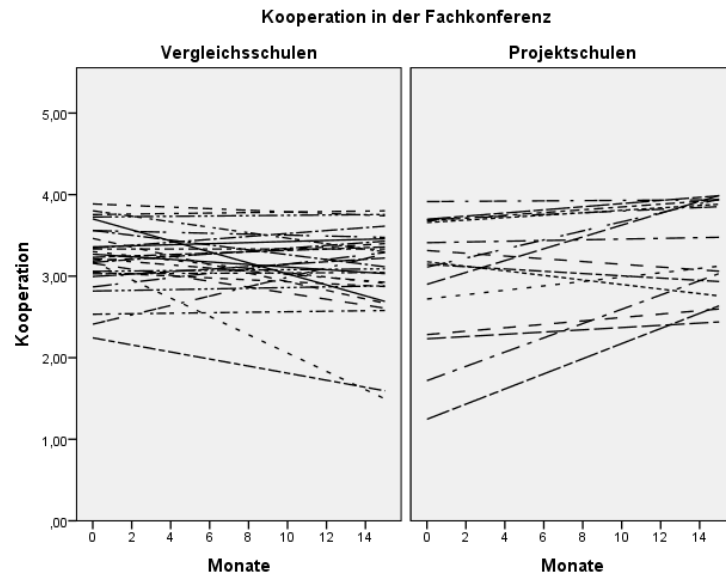


Abbildung 23. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal Kooperation in der Fachkonferenz.

Die Analysen im Mehrebenenmodell bestätigen die bisherigen Ergebnisse und ermöglichen eine differenzierte Analyse, in der eine Gruppenunterscheidung möglich ist und die unterschiedlichen Testzeitpunkte in Rechnung gestellt werden können.

Für das Merkmal der Zusammenarbeit in der Fachkonferenz zeigt sich, dass die Lehrkräfte in den Projektschulen zum Beginn des Projektes signifikant weniger zusammenarbeiten als ihre Kolleginnen und Kollegen in den Vergleichsschulen. Der durchschnittliche Skalenwert einer Lehrkraft aus den Vergleichsschulen liegt bei 2.5 Punkten. Die Lehrkräfte in den Projektschulen erreichen im Mittel nur 2.2 Skalenpunkte. Allerdings steigert sich die Zusammenarbeit in der Fachkonferenz bei den Lehrkräften in den Vergleichsschulen signifikant bis zum Projektende, wobei die Lehrkräfte in den Projektschulen trotz geringerer Ausgangslage keine bedeutende Entwicklung verzeichnen. Das höhere Wachstum pro Monat um 0.003 Skalenpunkte in den Projektschulen ist somit nicht systematisch (Tabelle 37).

Hinsichtlich der Kooperation in der Fachkonferenz ergeben sich andere Ergebnisse. Die Lehrkräfte in den Vergleichsschulen kooperieren im Mittel zum Beginn des Projektes in Höhe von 3.1 Skalenpunkten. Eine durchschnittliche Lehrkraft in den Projektschulen kooperiert nur unwesentlich weniger (0.02 Skalenpunkte) mit den Kollegen in der Fachkonferenz. Die Gruppen unterscheiden sich demnach in ihrer Ausgangssituation nicht.

In den Vergleichsschulen, die keine unterstützende Begleitung bei der Entwicklung eines kompetenzorientierten Unterrichts erhielten, veränderte sich die Kooperation auch nicht.

Tabelle 37. Wachstumseffekte (Modell C) der Zusammenarbeit und Kooperation in der Fachkonferenz, Studie 1, Lehrkräfte

		Parameter	Zusammenarbeit	Kooperation
<i>Feste Effekte</i>				
Ausgangswert	Konstante	γ_{00}	2.499*** (.056)	3.072*** (.073)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	-.265* (.110)	-.024 (.147)
Wachstumswert	Konstante	γ_{10}	.010* (.005)	-.007 (.006)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	.013 (.008)	.021* (.010)
<i>Zufallseffekte</i>				
	innerhalb Personen	σ^2_e	.095	.050
Level 1	Ausgangswert	σ^2_0	.096***	.250***
Level 2	Wachstumswert	σ^2_1	.000	.001***

Anmerkung. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

Das mit 0.007 Skalenpunkten quantifizierte monatliche Sinken ist aufgrund der fehlenden Signifikanz als zufällig zu betrachten. Die Lehrkräfte in den Projektschulen, die mit Projektbeginn an einer Weiterentwicklung des Mathematikunterrichts gemeinsam arbeiteten, steigerten erwartungsgemäß ihre Kooperation in der Fachkonferenz signifikant um monatlich 0.02 Skalenpunkte.

6.2.1.3 Die mathematischen Kompetenzen

Die mit einer Umstellung auf einen kompetenzorientierten Unterricht auszubildenden mathematischen Kompetenzen wurden ebenfalls einer statistischen Prüfung unterzogen. Gerade mit Blick auf den inhaltlichen Input durch die in den Projektschulen realisierten Treffen und Beratungen, sind hier große Unterschiede zwischen den Lehrkräften zu erkennen, die durchgängig am Projekt teilnahmen (Längsschnitt) und denen, die die Teilnahme nur phasenweise bescheinigten (Abbildung 24 ff.). Die Daten zeigen, dass die Lehrkräfte des Längsschnitts ihre kompetenzorientierten Tätigkeiten im Mathematikunterricht stark veränderten und damit vom Projekt intensiv beeinflusst wurden.

Tabelle 38. Deskriptive Statistiken zum Merkmal des mathematischen Argumentierens

Merkmal	MZP 1				MZP 2				MZP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
mathematisches Argumentieren (Querschnitt)												
Projektgruppe	52	4.08 (0.87)			27	4.15 (0.74)			25	3.97 (0.71)		
Vergleichsgruppe	60	4.05 (0.80)	0.05	.00	55	4.07 (0.69)	0.20	.00	48	3.91 (0.76)	0.12	.00
gesamt	112	4.06 (0.83)			82	4.10 (0.70)			73	3.93 (0.74)		
mathematisches Argumentieren (Längsschnitt)												
Projektgruppe	16	3.86 (0.96)			16	4.29 (0.76)			16	4.04 (0.73)		
Vergleichsgruppe	28	3.91 (0.83)	0.04	.00	28	3.97 (0.72)	1.89	.04	28	3.72 (0.71)	2.00	.05
gesamt	44	3.90 (0.87)			44	4.09 (0.74)			42	3.84 (0.73)		
mathematisches Argumentieren (imputiert)												
Projektgruppe	26	3.98 (0.90)			26	4.15 (0.74)			26	4.01 (0.62)		
Vergleichsgruppe	46	3.89 (0.72)	0.21	.00	46	4.07 (0.72)	0.18	.00	46	3.85 (0.71)	1.04	.02
gesamt	72	3.93 (0.79)			72	4.10 (0.72)			72	3.91 (0.68)		

Anmerkungen. $\sim p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Der Unterricht in Bezug auf das mathematische Argumentieren scheint in allen teilnehmenden Schulen ähnlich ausgeprägt zu sein. Von den nahezu gleichen Ausgangsniveaus springen die Lehrkräfte in den Projektschulen jedoch auf ein 0.4 Skalenpunkte erhöhtes Niveau (Längsschnitt), was sich im fortlaufenden Projekt wieder verliert, jedoch nicht auf das Anfangsniveau zurückfällt. Eine ähnliche Entwicklung, wenn auch nicht so stark ausgeprägt ist auch für die Vergleichsschulen zu beobachten. Auch hier scheint es zunächst eine Steigerung zu geben, die sich zum dritten Messzeitpunkt jedoch wieder verliert und sogar unter das Ausgangsniveau abrutscht. Damit argumentieren die Schülerinnen und Schüler in den Vergleichsschulen am Ende der Projektlaufzeit in den Vergleichsschulen weniger als zu Beginn des Projektes. Gründe für den Rückgang dieser mathematischen Kompetenz könnten im schulinternen Stoffverteilungsplan liegen (da alle Schulen davon betroffen sind) oder in der Erkenntnis der Lehrkräfte, dass einige Teilgebiete der Mathematik sich nur bedingt zum mathematischen Argumentieren eignen. Dennoch scheinen die durchgängig im Projekt arbeitenden Lehrkräfte die Kompetenz besser in den Unterricht integrieren zu können als die Lehrkräfte in den Vergleichsschulen.

Eine Varianzanalyse mit Messwiederholung auf einem Faktor ergab einen Haupteffekt für das Merkmal über die Zeit ($F(2, 84) = 4.73$; $p = .011$). Zusätzlich entwickeln sich auch die Gruppen über die Zeit bedeutend unterschiedlich ($F(2, 84) = 2.69$; $p = .074$). Die

Lehrkräfte der Projektgruppen konnten in Bezug auf dieses Merkmal eine signifikante Verbesserung erreichen im Vergleich zu den Lehrkräften in den Vergleichsschulen.

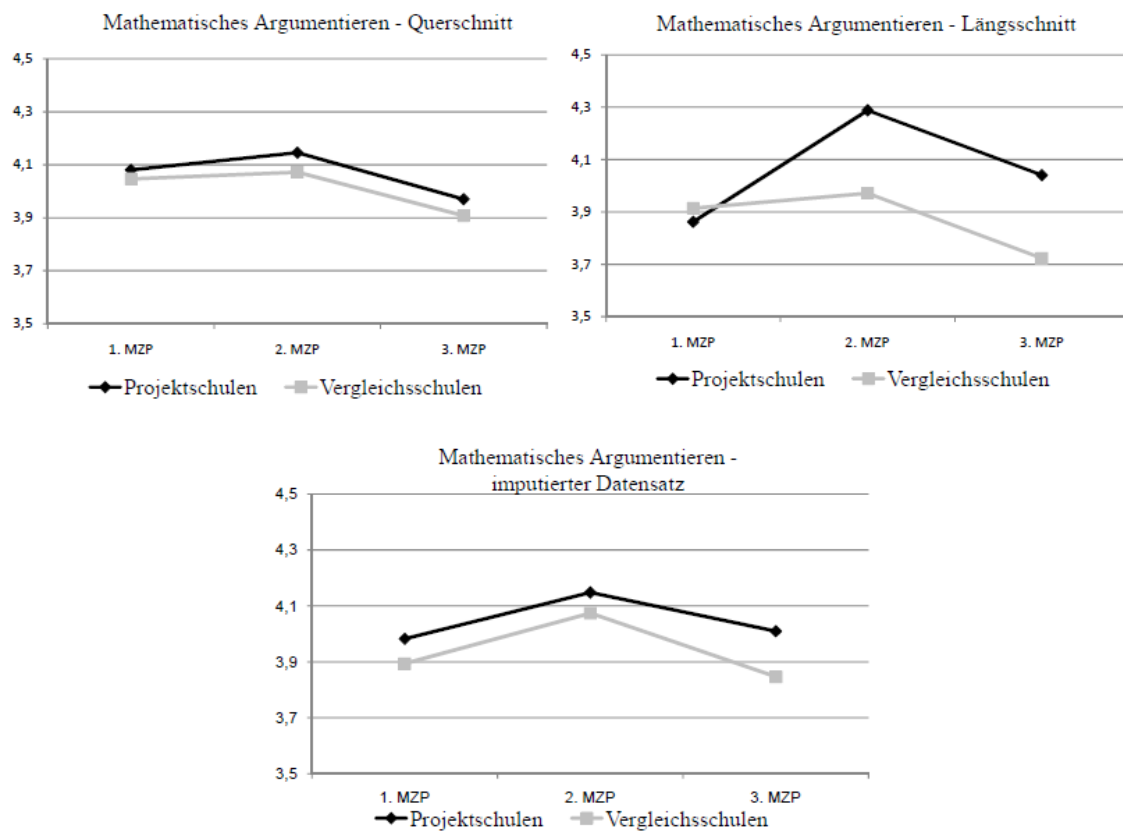


Abbildung 24. Mathematisches Argumentieren im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

In der intraindividuellen Betrachtung der geglätteten Regressionsgeraden zeigt sich zunächst eine höhere Variabilität in den Ausgangswerten für die Projektsschulen, die sich zum Ende deutlich verringert und im Mittel auf einem höheren Niveau endet. Die Aussage kann für die Lehrkräfte der Vergleichsschulen nicht getätigt werden. In Abbildung 25 sind neben drei erkennbaren Steigerungen auch drei erkennbare Reduzierungen in der Merkmalsausprägung des mathematischen Argumentierens erkennbar. Eine Entwicklung ist in den Vergleichsschulen nur schwer zu beobachten.

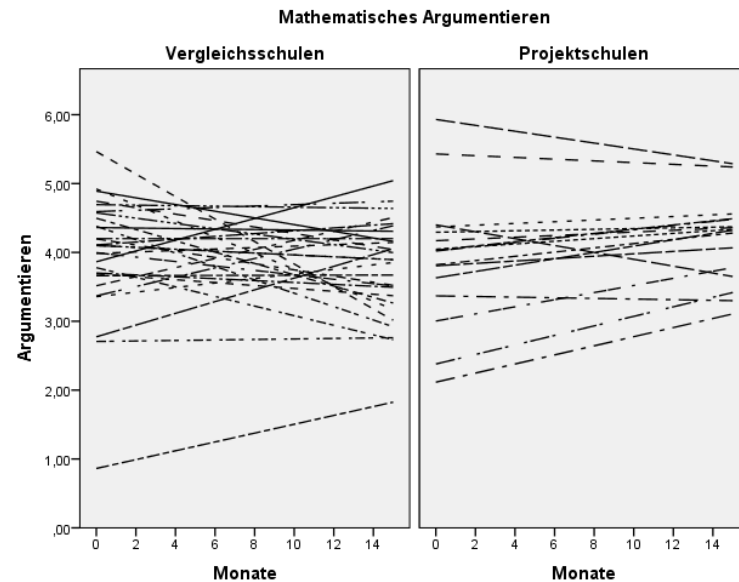


Abbildung 25. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisches Argumentieren.

Das Ausgangsniveau für die mathematische Kompetenz des Probleme Lösens zeigt große Unterschiede auf. Die Lehrkräfte des Längsschnitts unterscheiden sich in mehr als 0.4 Skalenpunkten voneinander (Abbildung 26), was sich statistisch durch einen signifikanten Mittelwertvergleich bestätigen lässt ($F(1, 42) = 4.523$; $p = .039$; $\eta^2 = .097$). Die Schulen scheinen hinsichtlich dieses Merkmals einen stark variierenden Unterricht anzubieten. Im Mittel geben die Lehrkräfte in den Vergleichsschulen an, viel öfter mathematische Probleme zu lösen als in den Projektschulen. Die Projektschulen stellen auch hier eine Negativauswahl dar.

Im Verlauf des Projektes initiieren die Lehrkräfte der Projektschulen viel häufiger Schülertätigkeiten, die mathematisches Probleme lösen erfordern. Eine weitere Ursache für den starken Anstieg könnte auch in dem Verständnis der Kompetenz liegen. Die Mitarbeit im Projekt könnte dazu geführt haben, diese Kompetenz inhaltlich besser definieren zu können. Für die Vergleichsschulen ist eine Abnahme in der Initiierung solcher Tätigkeiten zu beobachten. Gründe hierfür könnten im mathematischen Teilgebiet liegen, das zum Messzeitpunkt bearbeitet wurde, liegen.

Tabelle 39. Deskriptive Statistiken zum Merkmal Probleme mathematisch Lösen

Merkmal	MZP 1				MZP 2				MZP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
Probleme mathematisch Lösen (Querschnitt)												
Projektgruppe	52	4.21 (0.80)			27	4.37 (0.66)			25	4.30 (0.74)		
Vergleichsgruppe	59	4.38 (0.72)	1.44	.01	55	4.24 (0.79)	0.52	.01	48	4.20 (0.74)	0.33	.01
gesamt	111	4.30 (0.76)			82	4.28 (0.75)			73	4.24 (0.73)		
Probleme mathematisch Lösen (Längsschnitt)												
Projektgruppe	16	3.96 (0.65)			16	4.55 (0.64)			16	4.38 (0.73)		
Vergleichsgruppe	28	4.40 (0.66)	4.52*	.10	28	4.15 (0.82)	2.87~	.06	28	4.09 (0.77)	1.49	.03
gesamt	44	4.24 (0.68)			44	4.30 (0.78)			42	4.19 (0.76)		
Probleme mathematisch Lösen (imputiert)												
Projektgruppe	26	4.10 (0.70)			26	4.39 (0.67)			26	4.38 (0.62)		
Vergleichsgruppe	46	4.34 (0.63)	2.20	.03	46	4.25 (0.78)	0.58	.01	46	4.17 (0.74)	1.51	.02
gesamt	72	4.25 (0.66)			72	4.30 (0.74)			72	4.24 (0.71)		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Es ergibt sich ein Interaktionseffekt der Gruppen über die Zeit, der bereits bei der grafischen Betrachtung vermutet werden konnte. Während der Haupteffekt über die Zeit, also die Entwicklung des Merkmals im Allgemeinen keine bedeutenden Unterschiede zwischen den Messzeitpunkten ausweist, zeigt die Gruppenbetrachtung eine eindeutig unterschiedliche Entwicklung. Die Ergebnisse einer zweifaktoriellen Varianzanalyse mit Messwiederholung auf einem Faktor bestätigt die Interaktion zwischen der zeitlichen und gruppenspezifischen Entwicklung ($F(2, 84) = 11.89$; $p = .000$). Die Projektschulen entwickeln das Merkmal des Probleme mathematisch Lösen signifikant besser als die Lehrkräfte der Vergleichsschulen.

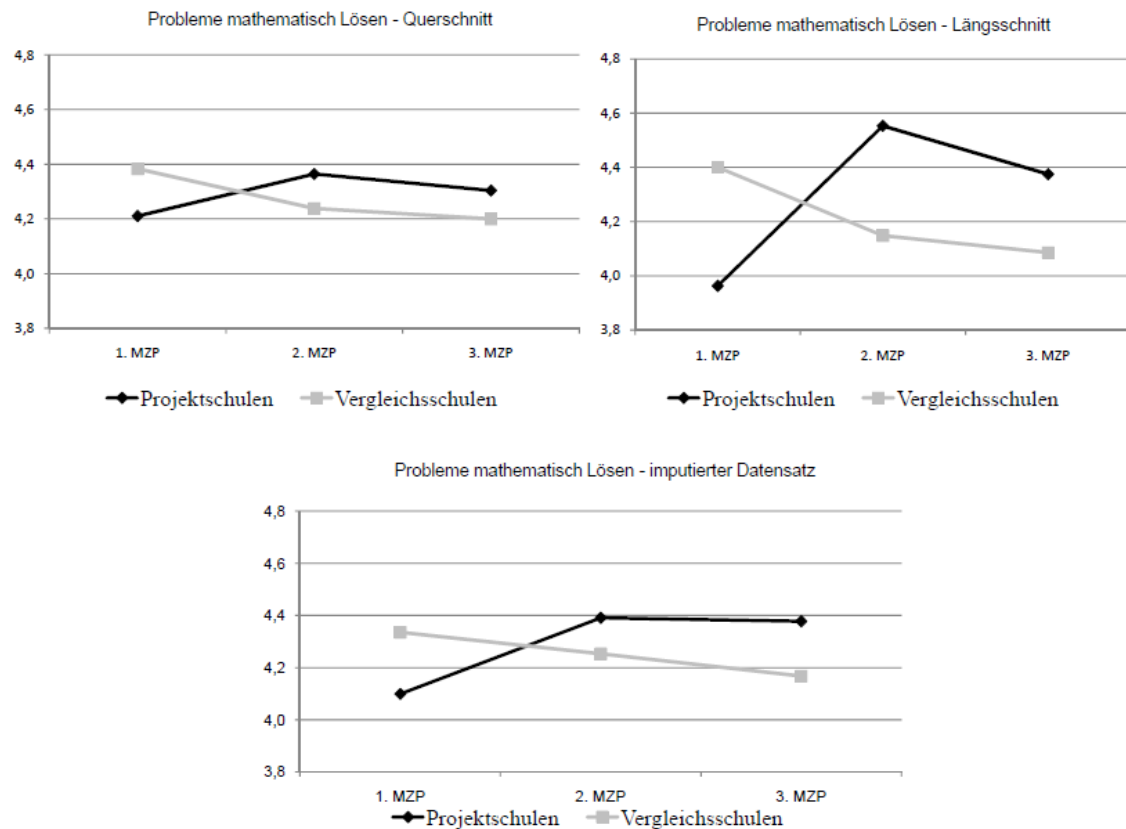


Abbildung 26: Probleme mathematisch Lösen im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

Die intraindividuellen Entwicklungen der einzelnen Lehrkräfte über die Projektlaufzeit weisen in den Vergleichsschulen einen Ausreißer mit sehr niedrigen Werten auf, dessen Häufigkeit der Initiierung mathematischer Problemlösungen sogar noch abnimmt. Ursache könnte hier eine abweichende Lehrplanarbeit sein. Insgesamt kann für die Vergleichsschulen ein leicht negativer Trend ausgemacht werden. In den Projektschulen überwiegt die gleichbleibende oder teilweise positive Entwicklung.

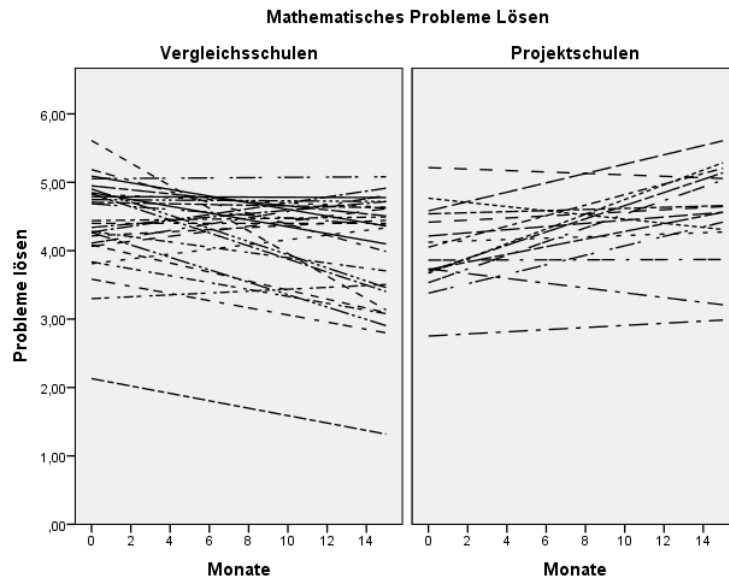


Abbildung 27. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal Probleme mathematisch Lösen.

Beide Gruppen weisen grafisch eine recht große Streuung auf, wodurch sich mit Blick auf die Projektschulen ein tendenziell positiver Trend ausmachen lässt, während die individuellen Entwicklungslinien der Vergleichsschulen in Abbildung 27 auf einen tendenziell negativen Verlauf hindeuten.

Im Merkmal des mathematischen Modellierens lassen sich trotz grafisch stark ausgeprägter Unterschiede in den Längsschnittdaten statistisch keine Gruppenunterschiede nachweisen, so dass von einheitlichen Anfangswerten auszugehen ist. In der Entwicklung ist ein sprunghafter Anstieg der Schülertätigkeit des mathematischen Modellierens bei den Lehrkräften des Längsschnittes erkennbar. Ein signifikanter Mittelwertunterschied zwischen den beiden Gruppen ist jedoch bei einzelner Betrachtung der Messzeitpunkte auch hier nicht nachweisbar.

Tabelle 40. Deskriptive Statistiken zum Merkmal mathematisches Modellieren

Merkmal	MZP 1				MZP 2				MZP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
mathematisches Modellieren (Querschnitt)												
Projektgruppe	52	3.90 (0.98)			27	4.09 (0.83)			25	4.16 (0.87)		
Vergleichsgruppe	59	3.98 (0.84)	0.19	.00	55	3.98 (0.90)	0.31	.00	48	4.03 (0.79)	0.39	.01
gesamt	111	3.94 (0.90)			82	4.02 (0.87)			73	4.08 (0.81)		
Mathematisches Modellieren (Längsschnitt)												
Projektgruppe	16	3.51 (0.97)			16	4.22 (0.95)			16	4.35 (0.78)		
Vergleichsgruppe	28	3.94 (0.85)	2.29	.05	28	3.90 (0.95)	1.21	.00	28	3.86 (0.78)	3.97~	.09
gesamt	44	3.78 (0.91)			44	4.01 (0.95)			42	4.04 (0.81)		
mathematisches Modellieren (imputiert)												
Projektgruppe	26	3.82 (0.96)			26	4.10 (0.79)			26	4.31 (0.67)		
Vergleichsgruppe	46	3.93 (0.78)	0.33	.01	46	4.01 (0.85)	0.21	.00	46	4.05 (0.73)	2.33	.03
gesamt	72	3.89 (0.84)			72	4.04 (0.83)			72	4.14 (0.72)		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Bei einer längsschnittlichen Betrachtung zeigt sich jedoch für die Lehrkräfte der Projekt-schulen, dass sie die Schülertätigkeiten des Modellierens vom Beginn bis zum Ende der Studie um mehr als 0.8 Skalenpunkte steigern konnten, während die Vergleichsschulen auf ihrem Anfangsniveau nahezu verharren (Längsschnitt). Ein Vergleich mit den Quer-schnittsdaten macht den Mehrwert der kontinuierlichen Projektarbeit deutlich. Es schei-nen nur die Lehrkräfte ihren Unterricht umzustellen, die kontinuierlich im Projekt arbei-ten. Ein Synergieeffekt auf die Kollegen kann nicht beobachtet werden, da sich die Lehr-kräfte in den Vergleichsschulen in ähnlicher Weise entwickeln. Das mathematische Mo-dellieren stellt sich als Merkmal dar, das kontinuierlich weiterentwickelt wird, wenn es eine entsprechende Begleitung aus der Fachdidaktik erfährt.

Abgesichert wird die in der Abbildung 28 dargestellte Entwicklung durch die Varianzana-lyse. Es ergibt sich sowohl ein signifikanter Haupteffekt über die Zeit ($F(2, 84) = 8.94$; $p = .000$) als auch ein signifikanter Interaktionseffekt der Gruppe*Zeit ($F(2, 84) = 12.20$; $p = .000$). Damit initiieren die Lehrkräfte der Projektschulen über die Projektlaufzeit sig-nifikant häufiger Tätigkeiten, die mathematisches Modellieren zu Grunde legen, als die Lehrkräfte in den Vergleichsschulen.

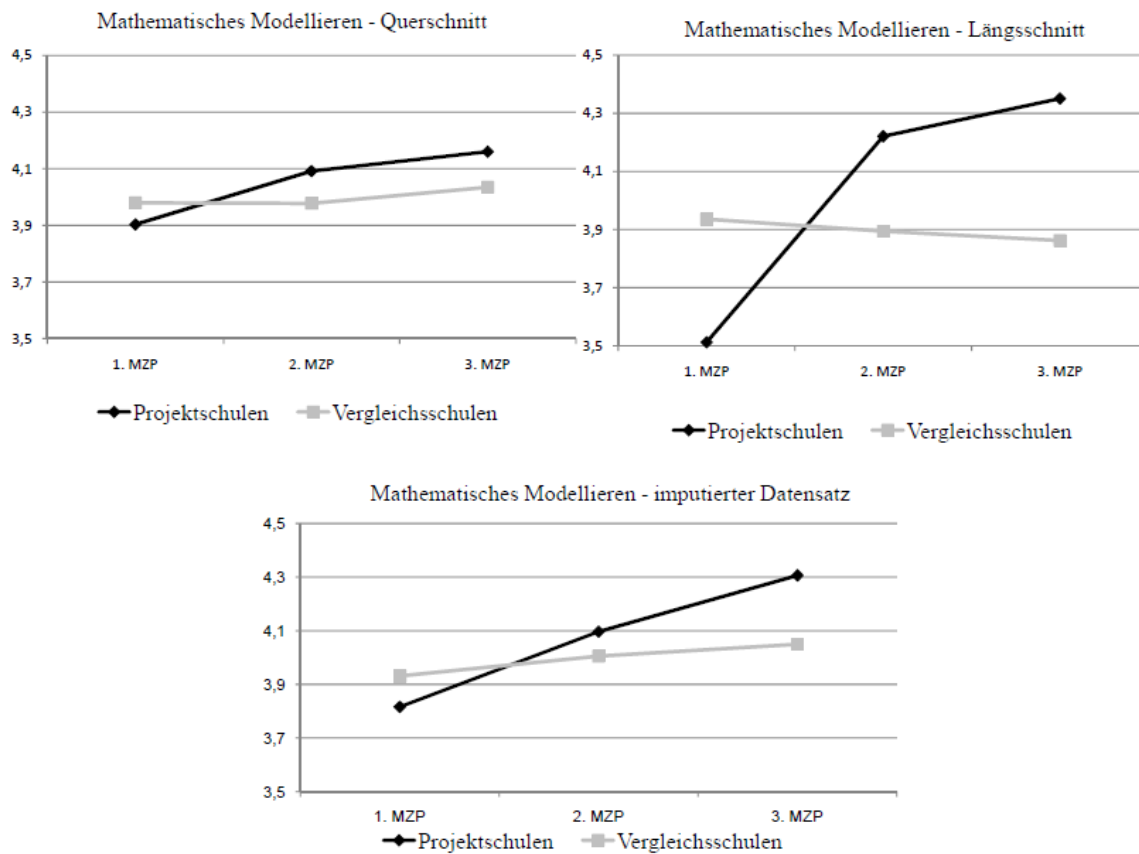


Abbildung 28. Mathematisches Modellieren im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

Dieses Ergebnis wird auch bei der intraindividuellen Betrachtung deutlich. Mit nur wenigen Ausnahmen weisen die Lehrkräfte der Projektschulen eine positive Entwicklung auf, deren Streuung sich zum Ende der Projektlaufzeit verringert. Für zwei Lehrkräfte der Projektschulen können signifikante positive Entwicklungen nachgewiesen werden. Im Gegensatz dazu stellen sich die Lehrkräfte der Vergleichsschulen als indifferent mit sowohl positiven als auch negativen Entwicklungen dar. Auch die Heterogenität der Daten nimmt zum Ende des Projekts nicht ab. In den Projektschulen kann davon ausgegangen werden, dass die inhaltliche Arbeit mit den kompetenzorientierten Aufgaben zu mehr Handlungssicherheit bei der mathematischen Kompetenzorientierung führt.

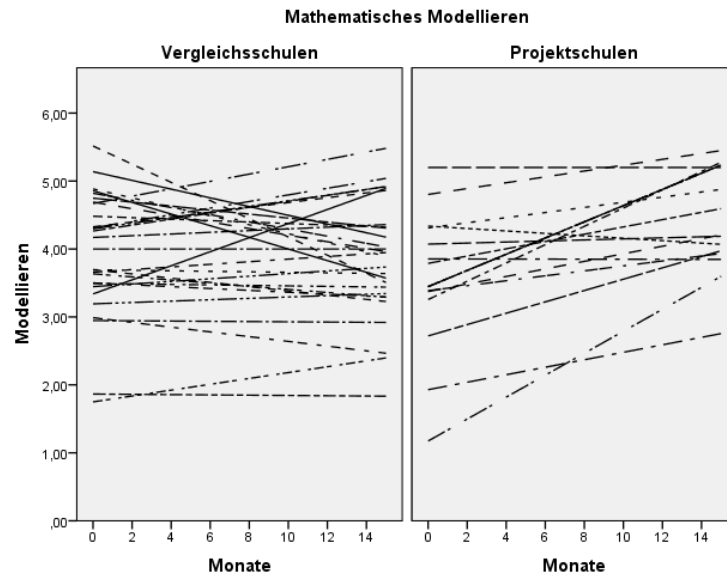


Abbildung 29. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisches Modellieren.

Die Kompetenz des Verwendens mathematischer Darstellungen variiert bedingt durch den Stoffverteilungsplan in den Schulen. Dennoch kann davon ausgegangen werden, dass sich mathematische Themen zu großen Teilen durch Tabellen und Diagramme bearbeiten lassen. Die Daten weisen auf eine einheitliche Ausgangslage hin, womit eine gleich häufige Anwendung der Kompetenz impliziert werden kann. Ziel der Projektarbeit ist es, diese mathematische Kompetenz bei den Schülerinnen und Schülern weiter auszubauen und damit eine verstärkte Kompetenzorientierung im Mathematikunterricht zu initiieren.

Tabelle 41. Deskriptive Statistiken zum Merkmal mathematische Darstellungen Verwenden

Merkmal	MZP 1				MZP 2				MZP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
Darstellungen Verwenden (Querschnitt)												
Projektgruppe	52	3.89 (0.78)			27	4.13 (0.61)			25	3.97 (0.77)		
Vergleichsgruppe	59	3.91 (0.74)	0.01	.00	55	4.00 (0.79)	0.53	.01	48	4.00 (0.71)	0.03	.00
gesamt	111	3.90 (0.76)			82	4.05 (0.74)			73	3.99 (0.73)		
Darstellungen Verwenden (Längsschnitt)												
Projektgruppe	16	3.81 (0.92)			16	4.28 (0.67)			16	4.15 (0.70)		
Vergleichsgruppe	28	3.87 (0.73)	0.06	.00	28	3.98 (0.78)	1.65	.04	28	3.86 (0.73)	1.60	.04
gesamt	44	3.85 (0.80)			44	4.08 (0.75)			42	3.97 (0.73)		
Darstellungen Verwenden (imputiert)												
Projektgruppe	26	3.91 (0.83)			26	4.17 (0.61)			26	4.08 (0.63)		
Vergleichsgruppe	46	3.84 (0.66)	0.24	.00	46	4.05 (0.78)	0.50	.01	46	3.93 (0.69)	0.85	.01
gesamt	72	3.86 (0.72)			72	4.09 (0.72)			72	3.98 (0.67)		

Anmerkungen. $\sim p < .10$; $*p < .05$; $**p < .01$; $***p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Im Verlauf der Projektarbeit zeigt sich jedoch, dass ein Unterschied zwischen den Schulen der Vergleichs- und der Projektgruppe nicht beobachtet werden kann. Während die Lehrkräfte des Längsschnitts durch die Projektarbeit dazu veranlasst werden, diese Schülertätigkeiten um nahezu 0.5 Skalenpunkte häufiger zu initiieren, verzerren ihre Kollegen, die nur unregelmäßig am Projekt teilnehmen diesen Wert für die gesamte Stichprobe nach unten, sodass hier nur noch ein Wachstum von ca. 0.3 Skalenpunkten zu erkennen ist (Vergleich 1. und 2. Messzeitpunkt, Querschnitt und Längsschnitt). Dieser Wert sinkt im 3. Messzeitpunkt im Querschnitt fast wieder auf das Anfangsniveau zurück (Querschnitt), während die durchgängige Projektarbeit (Längsschnitt) dazu geführt hat, dass die Lehrkräfte auch weiterhin diese Kompetenz in ihren Unterricht aufnehmen. Dennoch können die grafisch sichtbaren Unterschiede zwischen den beiden Gruppen statistisch nicht bestätigt werden ($F(2, 84) = 2.19; p > .05$). Es ergibt sich jedoch ein Haupteffekt über die Zeit, womit eine bedeutende Veränderung des Merkmals als Ganzes nachgewiesen werden kann ($F(2, 84) = 4.22; p = .018$).

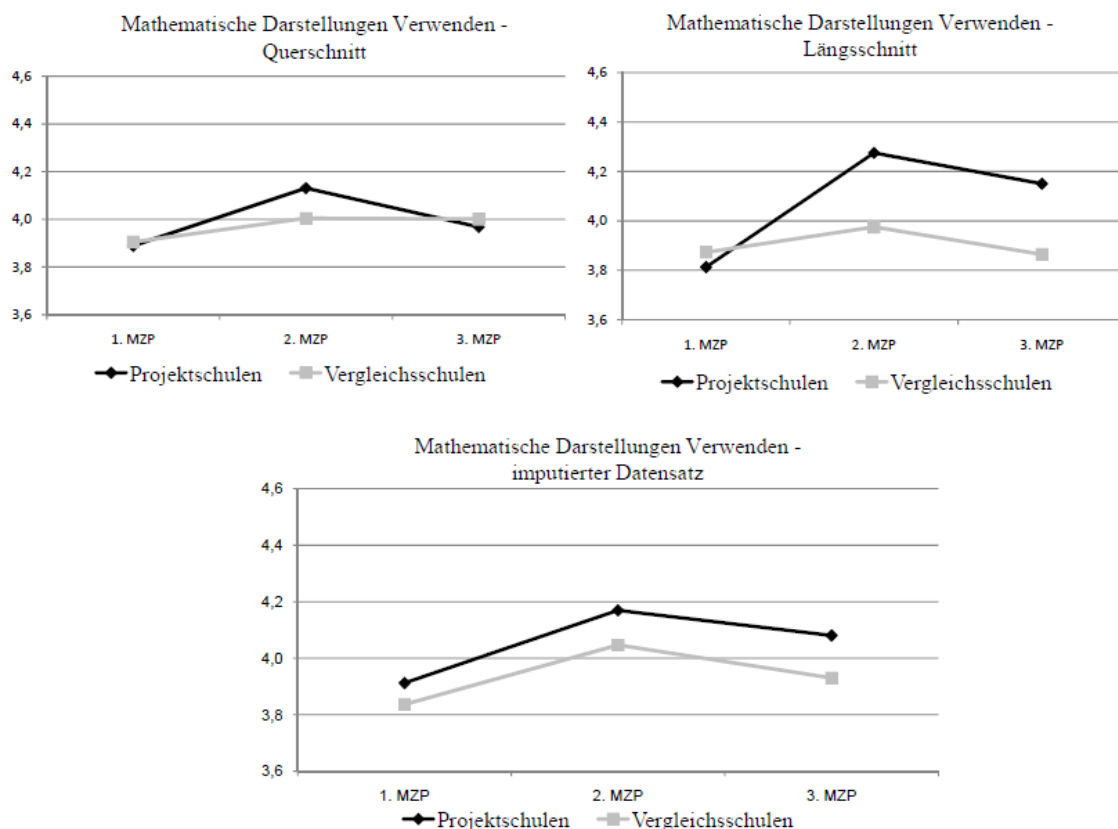


Abbildung 30. Mathematische Darstellungen Verwenden im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

In der intraindividuellen Betrachtung kann für die Lehrkräfte der Projektsschulen in nahezu allen Fällen eine ansteigende Entwicklung beobachtet werden. Die Ausnahmen bilden zwei Lehrkräfte, die diese Tätigkeiten in ihrem Unterricht jedoch nur sehr gering zurück-

fahren. Die Lehrkräfte der Vergleichsschulen zeigen sehr viel deutlicher einen negativen Trend in der Initiierung von Tätigkeiten, die mit Darstellungen und Diagrammen arbeiten.

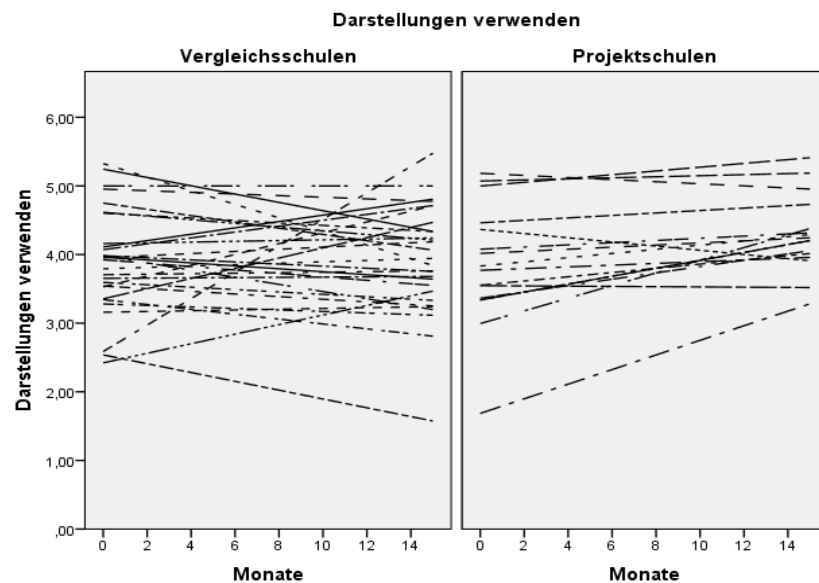


Abbildung 31. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematische Darstellungen Verwenden.

Die bei deutschen Schülerinnen und Schülern stark ausgeprägte Kompetenz des technischen Arbeitens wird auch von Seiten der Lehrkräfte in dieser Studie bestätigt. Die Unterrichtsgestaltung der Lehrkräfte verweist auf eine von Beginn an häufige Aufforderung zur Abarbeitung von mathematischen Routinen. Diese mathematische Kompetenz wird auch im Verlauf des Projektes nicht zu Gunsten einer anderen mathematischen Kompetenz abgebaut. Vielmehr wird sie über beide Schuljahre hinweg im Mittel nahezu wöchentlich im Unterricht initiiert.

Projekt- und Vergleichsschulen unterscheiden sich in diesem Merkmal nicht signifikant voneinander. Die Effektstärken zur Einordnung der Gruppenunterscheidung innerhalb der einzelnen Messzeitpunkte liegen fast ausnahmslos bei einem Wert von Null. Damit zeigt sich in diesem Merkmal eine gewisse Homogenität des Mathematikunterrichts.

Ergebnisse

Tabelle 42. Deskriptive Statistiken zum Merkmal *technisch Arbeiten*

Merkmal	MZIP 1				MZIP 2				MZIP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
technisch Arbeiten (Querschnitt)												
Projektgruppe	52	4.54 (0.57)			27	4.63 (0.49)			25	4.58 (0.64)		
Vergleichsgruppe	59	4.65 (0.67)	0.82	.01	55	4.62 (0.64)	0.01	.00	48	4.63 (0.61)	0.11	.00
gesamt	111	4.60 (0.63)			82	4.63 (0.59)			73	4.61 (0.62)		
technisch Arbeiten (Längsschnitt)												
Projektgruppe	16	4.70 (0.48)			16	4.74 (0.44)			16	4.72 (0.46)		
Vergleichsgruppe	28	4.72 (0.71)	0.01	.00	28	4.69 (0.74)	0.05	.00	28	4.71 (0.63)	0.00	.00
gesamt	44	4.71 (0.63)			44	4.71 (0.64)			42	4.71 (0.57)		
technisch Arbeiten (imputiert)												
Projektgruppe	26	4.68 (0.48)			26	4.64 (0.45)			26	4.77 (0.43)		
Vergleichsgruppe	46	4.61 (0.63)	0.27	.00	46	4.66 (0.65)	0.04	.00	46	4.68 (0.55)	0.68	.01
gesamt	72	0.64 (0.58)			72	4.66 (0.58)			72	4.71 (0.51)		

Anmerkungen. ~p<.10; *p<.05; **p<.01; ***p<.001, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

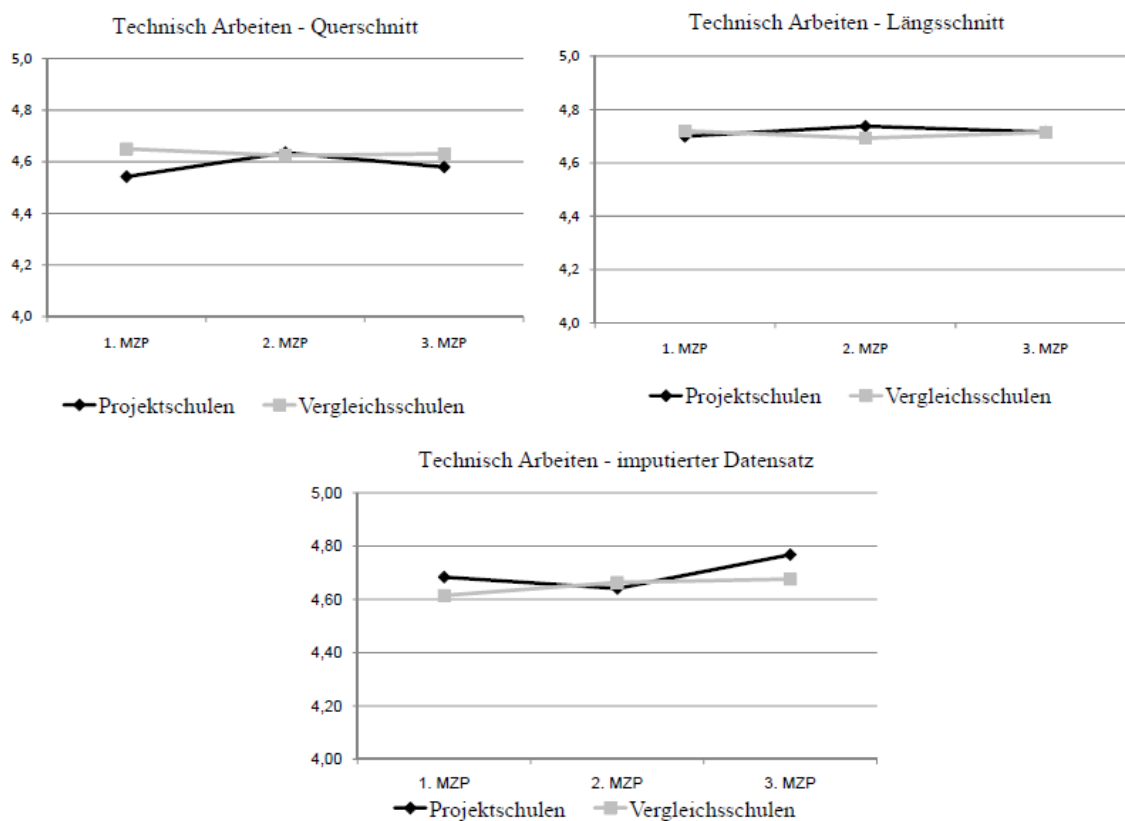


Abbildung 32. Technisch Arbeiten im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.

Ein Blick in die intraindividuellen Entwicklungen weist auf ein heterogenes Bild hin. Es sind sowohl in den Projekt- als auch den Vergleichsschulen einige Lehrkräfte auszumachen, die eine Verringerung von routinierten Abarbeitungstätigkeiten in ihrem Unterricht planten. Auf der anderen Seite wurde diese Kompetenz jedoch auch noch weiter ausgebaut, auch von einem bereits relativ hohen Ausgangsniveau. Es zeigt sich jedoch, dass die Steigungen bei den Lehrkräften in den Projektschulen flacher und insgesamt auf einem niedrigeren Niveau durchgeführt werden.

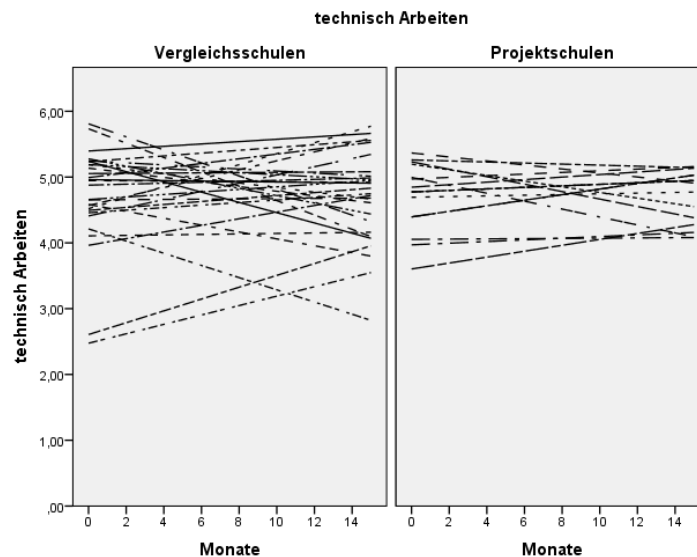


Abbildung 33. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal technisch Arbeiten.

Im mathematischen Kommunizieren sind in beiden Gruppen keine Entwicklungen zu beobachten. Ausgehend vom einem hohen Anfangswert steigern die Projektschulen die Tätigkeiten der Kommunikationskompetenz jedoch mehr als die Vergleichsschulen, fallen im Verlauf des Projektes jedoch wieder auf den Anfangswert zurück (Querschnitt).

Tabelle 43. Deskriptive Statistiken zum Merkmal *mathematisch Kommunizieren*

Merkmal	MZIP 1				MZIP 2				MZIP 3			
	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2	N	MW (SD)	F	η^2
mathematisch Kommunizieren (Querschnitt)												
Projektgruppe	52	4.59 (0.83)			27	4.80 (0.93)			25	4.59 (0.69)		
Vergleichsgruppe	59	4.59 (0.77)	0.00	.00	55	4.74 (0.76)	0.11	.00	48	4.67 (0.64)	0.24	.00
gesamt	111	4.59 (0.80)			82	4.76 (0.81)			73	4.64 (0.65)		
mathematisch Kommunizieren (Längsschnitt)												
Projektgruppe	16	4.45 (0.97)			16	4.89 (1.04)			16	4.70 (0.64)		
Vergleichsgruppe	28	4.66 (0.76)	0.65	.02	28	4.70 (0.83)	0.43	.01	28	4.67 (0.50)	0.03	.00
gesamt	44	4.58 (0.83)			44	4.77 (0.90)			42	4.68 (0.55)		
mathematisch Kommunizieren (imputiert)												
Projektgruppe	26	4.60 (0.85)			26	4.80 (0.95)			26	4.70 (0.49)		
Vergleichsgruppe	46	4.60 (0.69)	0.05	.00	46	4.77 (0.81)	0.03	.00	46	4.63 (0.57)	0.28	.00
gesamt	72	4.60 (0.72)			72	4.78 (0.83)			72	4.65 (0.56)		

Anmerkungen. $\sim p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Die Lehrkräfte in den Projektschulen, die durchgängig am Projekt teilnahmen (Längsschnitt), zeigen im Verlauf jedoch eine deutlichere Steigerung und verbleiben am Ende auch auf einem höheren Niveau als zu Beginn des Projektes. Diese Kompetenz scheint für die Lehrkräfte im Verlauf des Projektes an Bedeutung zugenommen zu haben, während die Lehrkräfte in den Vergleichsschulen keine Veränderung in ihrem Unterricht zeigen. Dennoch unterscheiden sich die Gruppen innerhalb der Messzeitpunkte statistisch nicht voneinander.

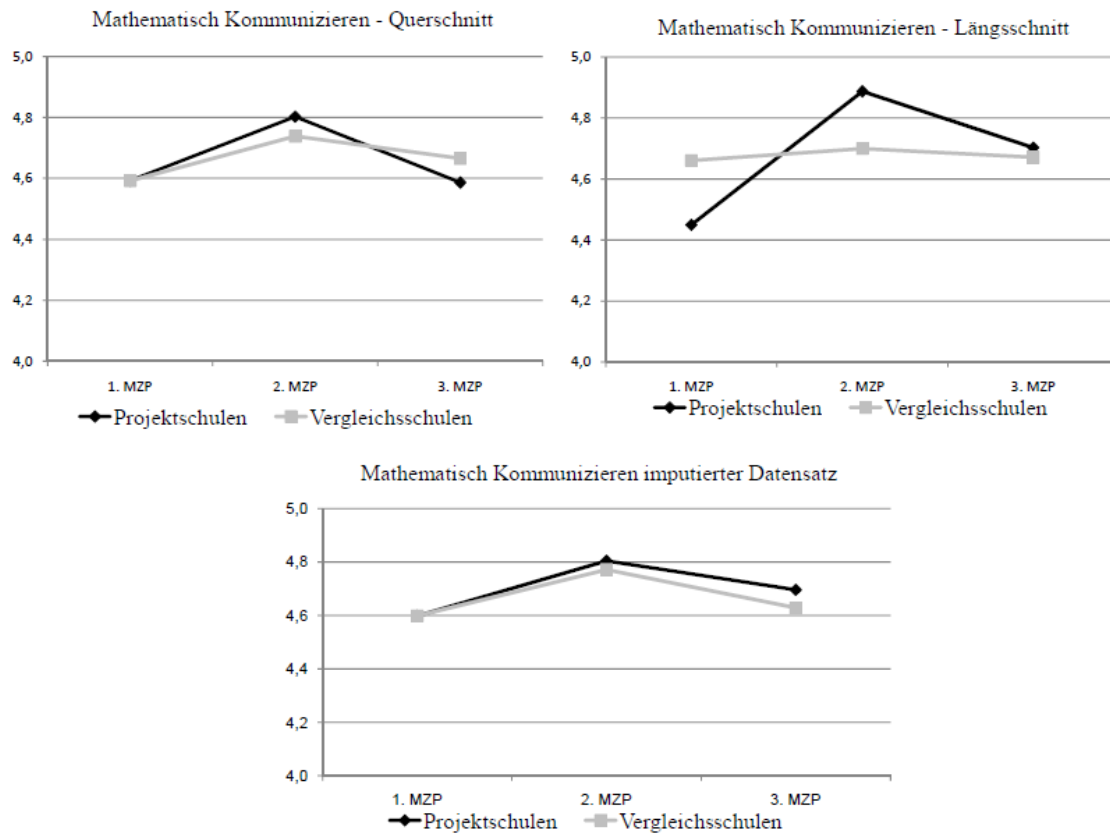


Abbildung 34. Mathematisches Kommunizieren im Verlauf, Gegenüberstellung von Querschnitt, Längsschnitt und imputierten Daten.

Auf individueller Ebene ist eine hohe Streuung der Werte erkennbar. Während in den Projektschulen nur wenige Lehrkräfte eine rückläufige Häufigkeit in der Unterrichtsgestaltung hinsichtlich dieses Merkmals berichten, kann in den Vergleichsschulen keine Tendenz ausgemacht werden.

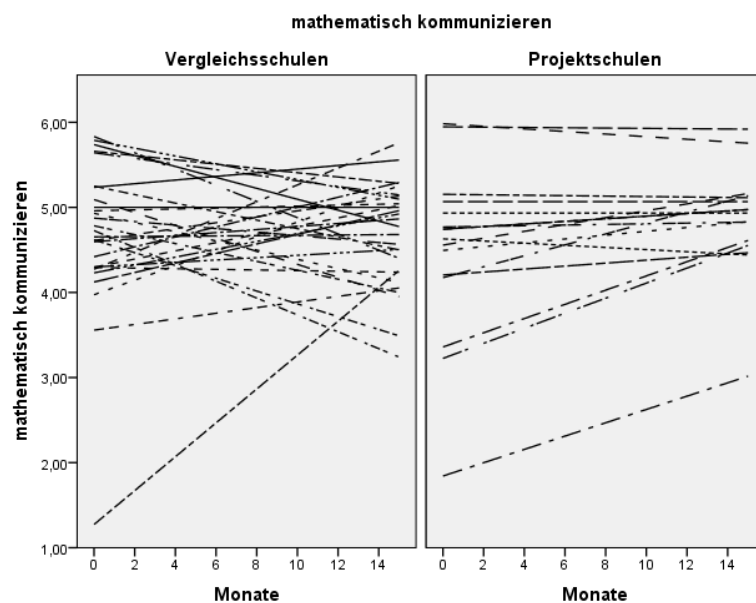


Abbildung 35. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisch kommunizieren.

Die grafisch vor allem im Vergleich zwischen Quer- und Längsschnitt sichtbaren Veränderungen in der Unterrichtsgestaltung werden durch die Analyse mit Hilfe von Mehrebenenmodellen relativiert. Einzig im Bereich des mathematischen Probleme lösens kann die signifikante Steigerung dieser kompetenzorientierten Tätigkeit für die Lehrkräfte in den Projektschulen gehalten werden. Während die Lehrkräfte in den Vergleichsschulen diese Kompetenz um monatlich 0.013 Skalenpunkte abbauten, wurde in den Projektschulen von einem anfänglich niedrigeren Ausgangsniveau eine Steigerung von 0.035 Skalenpunkten monatlich erreicht. Grundsätzlich geben die Lehrkräfte in den Projektschulen an, alle mathematischen Kompetenzen häufiger zu initiieren. Einzig im Bereich des technischen Arbeitens wurde auf eine weitere Intensivierung verzichtet. Die größten Veränderungen werden im Bereich des Probleme lösens und des Modellierens berichtet (Tabelle 44).

Die inhaltliche Begleitung der Lehrkräfte in der Phase der Implementation der Bildungsstandards und die damit verbundene Umsteuerung im Mathematikunterricht zeigt in den Projektschulen auf der Seite des Unterrichtsangebots tendenzielle Veränderungen, statistisch sind diese jedoch noch nicht bedeutend (Ausnahme: mathematische Probleme lösen). Die Ergebnisse zeigen jedoch, dass das Potential zur Umstellung des Mathematikunterrichts auf Kompetenzorientierung vorhanden und möglich ist, wenngleich sich das Unterrichtsangebot im Verlauf der Studie nicht auf die Abnehmerseite übertragen konnte (Abschnitt 6.3.1).

Ergebnisse

Tabelle 44. Wachstumseffekte (Modell C) der mathematischen Kompetenzen, Studie 1, Lehrkräfte

		Parameter	Argumentieren	Probleme Lösen	Modellieren	Darstellungen	technisch Arbeiten	Kommunizieren
<i>Feste Effekte</i>								
Ausgangswert	Konstante	γ_{00}	3.946*** (.114)	4.350*** (.102)	3.919*** (.126)	3.865*** (.105)	4.610*** (.099)	4.628*** (.114)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	.069 (.212)	-.233 (.167)	-.122 (.226)	.069 (.192)	.047 (.143)	-.012 (.212)
Wachstumswert	Konstante	γ_{10}	-.001 (.009)	-.013 (.010)	.010 (.010)	.010 (.009)	.006 (.008)	.005 (.010)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	.005 (.014)	.035* (.015)	.025 (.019)	.006 (.014)	.000 (.011)	.006 (.013)
<i>Zufallseffekte</i>								
Level 1	innerhalb Personen	σ^2_{ε}	.206	.204	.193	.236	.143	.252
Level 2	Ausgangswert	σ^2_0	.472***	.251***	.583***	.320***	.234***	.408***
	Wachstumswert	σ^2_1	.001*	.001**	.002***	.000	.000	.000

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

6.2.1.4 Zusammenfassung der Ergebnisse aus Studie 1

In Studie 1 wurden folgende Punkte besonders deutlich:

- Unterrichtsentwicklungskonzepte werden vom Interesse und den Einstellungen der Teilnehmer gegenüber dem Thema getragen. Lehrkräfte, die weniger kooperieren, den Bildungsstandards gegenüber kritischer eingestellt sind und die ihren eigenen Unterricht schlechter bewerten als ihre Kolleginnen und Kollegen, scheiden aus dem Projekt aus.
- Die Gegenüberstellung der Quer- und Längsschnittdaten zeigt, dass die Lehrpersonen im Längsschnitt die stärksten Veränderungen aufweisen. Eine sporadische Teilnahme an Unterrichtsentwicklungskonzepten zeigt nur eine geringe Wirkung auf die Veränderung der Unterrichtsgestaltung.
- Es zeigen sich keine Synergieeffekte im Kollegium im Hinblick auf die Wirkung der Multiplikatoren.
- Die gemeinsam entschiedene und durchgeführte Projektarbeit führt zu einer steigenden Kooperation in der Fachkonferenz.
- Die Lehrkräfte in den Projektschulen initiieren verstärkt kompetenzorientierte Tätigkeiten des mathematischen Probleme Lösen und verzichten auf einen Ausbau der Kompetenzen im technischen bzw. routinierten Arbeiten.

6.2.2 Studie 2 „Implementation der Bildungsstandards durch unterstützte Unterrichtsentwicklung“

In Studie 2 erfolgt eine Evaluation des Konzeptes des IQB zur Implementation der nationalen Bildungsstandards, indem hier eine zweite Vergleichsgruppe mit gleichem Studiendesign herangezogen wird. Die Vergleichsstichprobe aus dem Bundesland Hessen teilt sich ebenso in Projekt- und Vergleichsschulen auf. Das Kriterium zur Zuschreibung des Status Projektschule erfolgte nach der Teilnahme am Hessischen Projekt „Kompetenzorientiert Unterrichten“ (KOU). Im Folgenden werden die Projekt- und Vergleichsschulen grafisch gegenübergestellt. In Rechnung gestellt wird lediglich, ob eine Unterstützung bei der Unterrichtsentwicklung stattfand. Dabei werden die beiden Konzepte zunächst als gleichwertig betrachtet (globale Betrachtung). In einer differenzierten Darstellung erfolgt die Abbildung getrennt nach räumlichem und inhaltlichem Implementationskonzept. Den Ausführungen des Kapitels 5.4.5 für Studie 2 folgend, erfolgt die Darstellung hier nur für

die Lehrkräfte, die zu beiden Erhebungszeitpunkten an der Studie teilnahmen (Längsschnitt).

6.2.2.1 Einstellungen gegenüber der Einführung von nationalen Bildungsstandards

Bei der Betrachtung der positiven Einstellungen gegenüber den Bildungsstandards zeigt sich, dass diese nicht weiter ausgebaut werden konnten. Sowohl in den Projektschulen als auch den Vergleichsschulen konnte – global betrachtet – ein fast gleich starker Rückgang der Mittelwerte von .13 bzw. .17 Skalenpunkten beobachtet werden. Erwartet wurde eine positive Entwicklung der Lehrkräfte in den Projektschulen gegenüber den Vergleichsschulen.

Tabelle 45. Deskriptive Statistiken zum Merkmal positive Einstellungen gegenüber den Bildungsstandards

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>positive Einstellungen (global)</i>								
Projektgruppe	46	3.38 (0.11)			46	3.25 (0.10)		
Vergleichsgruppe	38	3.08 (0.11)	3.89	.05	38	2.91 (0.12)	4.96	.06
gesamt	84	3.24 (0.08)			84	3.10 (0.08)		
<i>positive Einstellungen (BB/BB)</i>								
Projektgruppe	12	3.87 (0.17)			12	3.74 (0.21)		
Vergleichsgruppe	17	3.32 (0.13)			17	3.36 (0.13)		
<i>positive Einstellungen (Hessen)</i>								
Projektgruppe	34	3.20 (0.12)	6.13	.19	34	3.08 (0.09)	10.92	.29
Vergleichsgruppe	21	2.88 (0.17)			21	2.54 (0.16)		
gesamt	84	3.24 (0.08)			72	3.10 (0.08)		

Anmerkungen. imputierter Datensatz; alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), *F*-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Bei einer differenzierten Betrachtung wird besonders grafisch sichtbar, dass sich die Gruppen dennoch unterschiedlich entwickeln. Auffällig ist zunächst, dass die Lehrkräfte in den Bundesländern BE/BB den Bildungsstandards grundsätzlich positiver gegenüberstehen als die Kolleginnen und Kollegen in Hessen. Die Gruppen unterscheiden sich signifikant voneinander, was durch die hohe Effektstärke qualifiziert wird. Da die Lehrkräfte in BE/BB das Projekt aus Studie 1 weiterführten, zeigen sich hier die Folgeeinstellungen der Lehrkräfte. Mit Verweis auf den dritten Messzeitpunkt in Studie 1 (Abschnitt 6.2.1.1) konnten die Lehrkräfte in den Projektschulen in BE/BB ihre positiven Einstellungen weiterhin ausbauen. Die Lehrkräfte in den Vergleichsschulen verharren auf dem in Studie 1

berichteten Niveau. Die Lehrkräfte aus Hessen mit ähnlichen Implementationskonzept drücken ihre positiven Einstellungen deutlich vorsichtiger aus. Trotz durchgeführtem Projekt liegen die Skalenwerte der positiven Einstellungen gegenüber den Bildungsstandards noch unter dem Niveau der Vergleichsschulen in BE/BB. Im 2. Messzeitpunkt zeigen einzig die Lehrkräfte in den Vergleichsschulen in Hessen, dass sie ihre nur gering ausgeprägten positiven Einstellungen gegenüber den Bildungsstandards noch weiter abbauen, während die verbleibenden Lehrkräfte auf dem Anfangsniveau verharren.

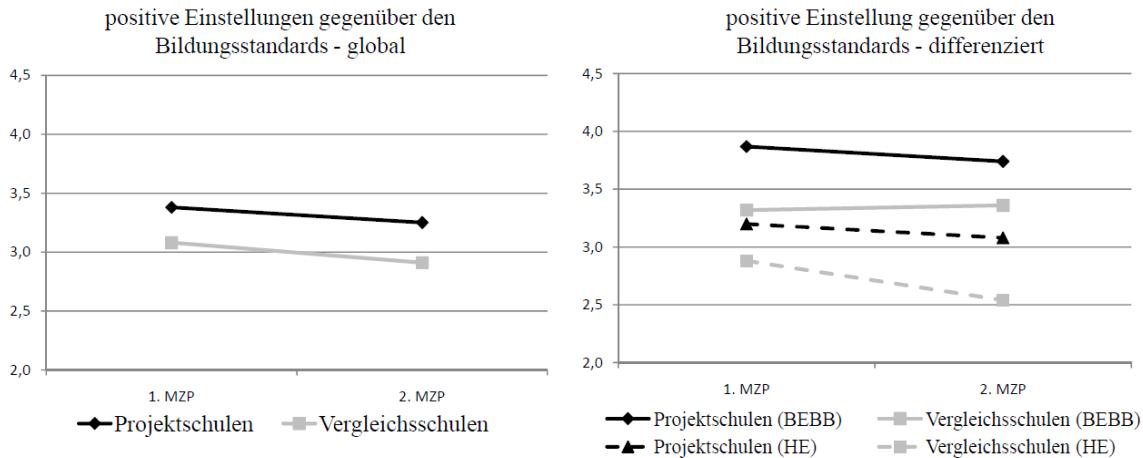


Abbildung 36. Positive Einstellungen gegenüber den Bildungsstandards - Veränderungen; Gegenüberstellung global und differenziert.

Eine Varianzanalyse mit Messwiederholung bestätigt den Gruppenunterschied auch statistisch. Es kann nur ein Haupteffekt für die Gruppe identifiziert werden ($F(3, 80)=11.47$, $p<.001$). Eine statistisch signifikante Veränderung über die Zeit zeigt sich nicht. Der Dunnett-T-Test mit der Referenzgruppe der Projektsschulen BE/BB zeigt, dass sich die Projektsschulen in BE/BB signifikant von den Projekt- und Vergleichsschulen in Hessen unterscheiden. Ein Vergleich der Differenzwerte der Projektsschulen BE/BB mit den Vergleichsschulen BE/BB ergab ein Signifikanz von $p=.062$.

Qualifiziert werden die Ergebnisse der positiven Einstellungen auch durch die Aussagen zu den negativen Einstellungen gegenüber den Bildungsstandards. Insgesamt stehen die Lehrkräfte in den Vergleichsschulen den Bildungsstandards sehr viel kritischer gegenüber als die Lehrkräfte in den Projektsschulen. Die beiden Gruppen unterscheiden sich in ihren Ausgangswerten ($F(1, 83)=5.03$, $p<.05$). Eine einfaktorielle Varianzanalyse mit Messwiederholung bestätigt die Unterschiedlichkeit der beiden Gruppen durch einen Haupteffekt ($F(1, 82)=6.10$, $p<.05$).

Tabelle 46. Deskriptive Statistiken zum Merkmal negative Einstellungen gegenüber den Bildungsstandards

Merkmal	MZIP 1				MZIP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>negative Einstellungen (global)</i>								
Projektgruppe	46	2.49 (0.11)			46	2.55 (0.11)		
Vergleichsgruppe	38	2.83 (0.11)	5.03	.06	38	2.85 (0.11)	3.59	.04
gesamt	84	2.64 (0.08)			84	2.69 (0.08)		
<i>negative Einstellungen (BB/BB)</i>								
Projektgruppe	12	2.02 (0.19)			12	2.03 (0.20)		
Vergleichsgruppe	17	2.61 (0.14)			17	2.57 (0.14)		
<i>negative Einstellungen (Hessen)</i>								
Projektgruppe	34	2.65 (0.12)	5.67	.18	34	2.74 (0.12)	6.53	.20
Vergleichsgruppe	21	3.01 (0.15)			21	3.08 (0.15)		
gesamt	84	2.64 (0.08)			84	2.69 (0.08)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In der differenzierten Darstellung ist erkennbar, dass die Lehrkräfte in Hessen stärker ausgeprägte negative Einstellungen aufweisen als die Lehrkräfte in BE/BB. In Hessen steigen die negativen Einstellungen sogar noch tendenziell an, während sie in BE/BB nahezu auf ihrem Anfangsniveau verharren. Der wünschenswerte Abbau der negativen Einstellungen ist in beiden Bundesländern und beiden Gruppen nicht eingetreten. Die Gruppen unterscheiden sich jedoch stark voneinander, was durch einen Haupteffekt der Gruppe statistisch nachgewiesen werden kann ($F(3, 80) = 7.60, p < .001$). Von der Referenzkategorie der Projektschulen (BE/BB) unterscheiden sich alle weiteren Gruppen signifikant voneinander (Dunnett-T-Tests).

Im Hinblick auf den evaluativen Aspekt des IQB-Fortbildungskonzeptes kann für die Einstellungen gegenüber den Bildungsstandards kein eindeutiges Ergebnis formuliert werden. Mit Blick auf die Vergleichsschulen des jeweiligen Bundeslandes zeigt sich, dass sich die Einstellungen der Lehrkräfte in nahezu gleichem Ausmaß verändern wie die der Lehrkräfte in den Projektschulen. Die Einstellungen einer Berufsgruppe gegenüber einer Innovation scheinen sich auch durch eine externe Unterstützung nur sehr träge zu verändern.

Ergebnisse

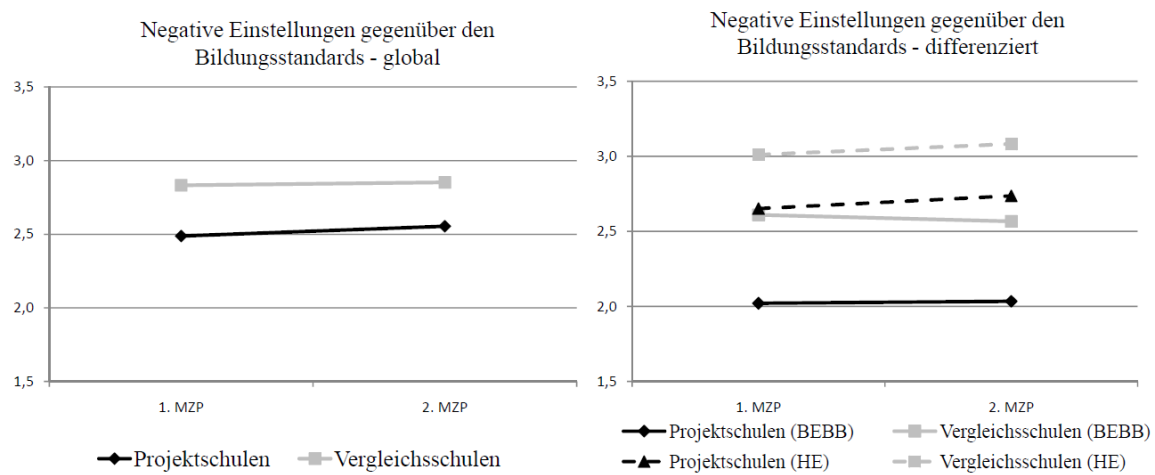


Abbildung 37. Negative Einstellungen gegenüber den Bildungsstandards - Veränderungen; Gegenüberstellung global und differenziert.

In den abschließenden Analysen zeigt sich, dass sich die Lehrergruppen in ihrem Ausgangsniveau der Studie 2 signifikant unterscheiden. Ein erwarteter signifikanter Anstieg der positiven Einstellungen und Abbau der negativen Einstellungen in den Projektschulen ist jedoch nicht zu beobachten.

Tabelle 47. Wachstumseffekte (Modell D) der Einstellungen gegenüber der Einführung von nationalen Bildungsstandards, Studie 2, Lehrkräfte

		Parameter	positive Einstellungen	negative Einstellungen
<i>Feste Effekte</i>				
Ausgangswert	Konstante	γ_{00}	3.405***	2.576***
			(.151)	(.144)
	Gruppe	γ_{01}	.393~	-.488*
(0=VS, 1=PS)			(.204)	(.186)
Bundesland		γ_{02}	-.482*	.436*
	(0=BE/BB, 1=HE)		(.197)	(.181)
Wachstumswert	Konstante	γ_{10}	-.008	-.005
			(.017)	(.014)
	Gruppe	γ_{11}	.008	.006
(0=VS, 1=PS)			(.023)	(.019)
Bundesland		γ_{12}	-.023	.016
	(0=BE/BB, 1=HE)		(.023)	(.018)
<i>Zufallseffekte</i>				
	innerhalb Personen	σ^2_{ε}	.223	.157
Level 1	Ausgangswert	σ^2_0	.174***	.265***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, BE/BB = Berlin/Brandenburg, HE = Hessen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparamters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

6.2.2.2 Kooperation und Zusammenarbeit in der Fachkonferenz

Bei der globalen Darstellung zwischen Projekt- und Vergleichsgruppe unterscheiden sich die Lehrkräfte in ihrem Anfangsniveau nicht voneinander. Grafisch ist jedoch ersichtlich, dass die Projektschulen um wenige Skalenpunkte stärker miteinander kooperieren als die Lehrkräfte in den Vergleichsschulen. Ein differenzierter Blick zeigt, dass die Lehrkräfte in BE/BB angeben, anfangs stärker miteinander zu kooperieren. Allerdings zeigt Tabelle 48 auch, dass diese Kooperation innerhalb eines Schuljahres um wenige Skalenpunkte abnimmt. Die Veränderung kann statistisch jedoch nicht nachgewiesen werden. Die differenzierte Darstellung zeigt weiterhin, dass die Lehrkräfte in den Projektschulen in BE/BB angeben, stark miteinander zu kooperieren. Ihre Werte liegen von Anfang an über denen der anderen Gruppen. Nimmt man diese Gruppe als Referenz für die Dunnett-T-Tests zur Gruppenunterscheidung, dann bestätigt sich jedoch nur eine signifikante Gruppenunterscheidung zu den Lehrkräften der Vergleichsschulen in Hessen. Der grafisch sichtbare Abstand zwischen den Projektschulen in BE/BB und den anderen Gruppen ist daher vorsichtig zu interpretieren.

Tabelle 48. Deskriptive Statistiken zum Merkmal Kooperation in der Fachkonferenz

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>Kooperation (global)</i>								
Projektgruppe	46	3.06 (0.09)			46	3.04 (0.09)		
Vergleichsgruppe	38	2.85 (0.11)	2.60	.03	38	2.82 (0.11)	2.28	.03
gesamt	84	2.97 (0.07)			84	2.94 (0.07)		
<i>Kooperation (BB/BB)</i>								
Projektgruppe	12	3.41 (0.15)			12	3.31 (0.24)		
Vergleichsgruppe	17	3.05 (0.14)			17	2.90 (0.13)		
<i>Kooperation (Hessen)</i>								
Projektgruppe	34	2.95 (0.10)	3.93	.13	34	2.94 (0.09)	1.94	.07
Vergleichsgruppe	21	2.68 (0.15)			21	2.76 (0.17)		
gesamt	84	2.97 (0.07)			84	2.94 (0.07)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Ebenfalls auffällig ist die einheitliche Entwicklung der befragten Lehrkräfte innerhalb der Bundesländer. Unabhängig von der unterstützenden Unterrichtsberatung (Intervention) entwickeln sich die Lehrkräfte in BE/BB sowohl in den Projekt- als auch den Vergleichsschulen leicht rückläufig und die Lehrkräfte in Hessen leicht positiv in Richtung stärkerer Kooperation. Beide Interventionen bzw. Konzepte scheinen hinsichtlich dieses Merkmals

Ergebnisse

keine Wirkung zu entfalten. Die Lehrkräfte aus BE/BB knüpfen an ihre Werte aus Studie 1 an, scheinen die Qualität der Zusammenarbeit jedoch nicht weiter ausbauen zu können.

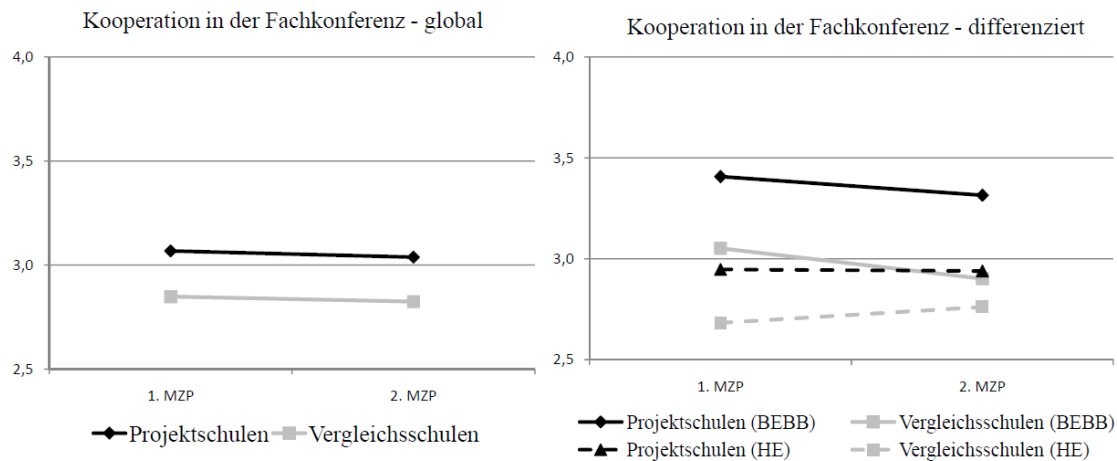


Abbildung 38. Kooperation in der Fachkonferenz - Veränderungen; Gegenüberstellung global und differenziert.

Überraschend sind daher die grafischen Ergebnisse zur Zusammenarbeit in der Fachkonferenz. Die Lehrkräfte in den Vergleichsschulen geben am Ende des Schuljahres an, häufiger in der Fachkonferenz zusammenzuarbeiten. Statistisch kann diese Entwicklung nicht nachgewiesen werden, wenngleich grafisch erkennbar ist, dass die Vergleichsschulen ihre Zusammenarbeit um mehr als 0.1 Skalenpunkte steigerten.

Tabelle 49. Deskriptive Statistiken zum Merkmal Zusammenarbeit in der Fachkonferenz

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>Zusammenarbeit (global)</i>								
Projektgruppe	46	2.57 (0.09)			46	2.55 (0.07)		
Vergleichsgruppe	38	2.40 (0.07)	2.19	.03	38	2.51 (0.10)	0.09	.00
gesamt	84	2.49 (0.06)			84	2.53 (0.06)		
<i>Zusammenarbeit (BB/BB)</i>								
Projektgruppe	12	2.75 (0.27)			12	2.60 (0.12)		
Vergleichsgruppe	17	2.54 (0.08)			17	2.55 (0.15)		
<i>Zusammenarbeit (Hessen)</i>								
Projektgruppe	34	2.50 (0.07)	2.18	.08	34	2.53 (0.08)	0.15	.01
Vergleichsgruppe	21	2.28 (0.11)			21	2.48 (0.14)		
gesamt	84	2.49 (0.06)			84	2.53 (0.06)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In der differenzierten Betrachtung wird deutlich, dass sich die Schulen zu Beginn des Schuljahres in der Zusammenarbeit noch stark unterscheiden. Zum Schuljahresende scheint es jedoch einen in allen Schulen ähnlichen Bedarf an Zusammenarbeit zu geben, da sich alle befragten Lehrkräfte zum zweiten Messzeitpunkt in ihren Antworten annähern. Offenbar kann der hohe Anfangswert in den Projektschulen BE/BB nicht aufrecht erhalten werden. Es scheint für dieses recht organisatorische Merkmal einen allgemeinen Konsens zu geben, wie viel Zusammenarbeit in den Fachkonferenzen nötig ist.

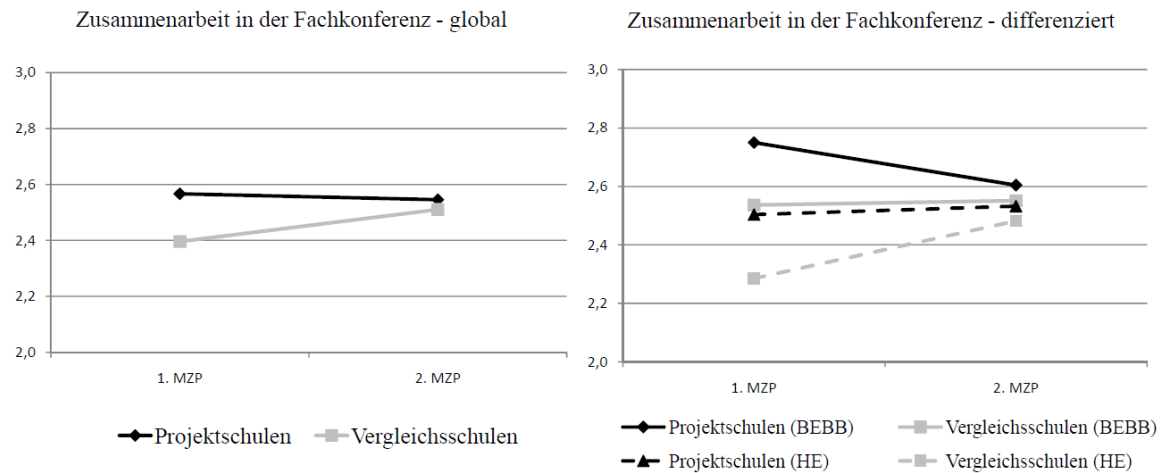


Abbildung 39. Zusammenarbeit in der Fachkonferenz - Veränderung; Gegenüberstellung global und differenziert.

Die Berechnungen im Mehrebenenmodell weisen ebenfalls die anfänglich starke Unterscheidung der Lehrergruppen nach.

Tabelle 51 zeigt für beide Skalen einen signifikant höheren Ausgangswert für die Projektschulen an. Weiterhin arbeiten und kooperieren die Schulen im Bundesland Hessen signifikant weniger in der Fachkonferenz als die Schulen in Berlin/Brandenburg. Die Wachstumsparameter der einzelnen Gruppen zeigen jedoch keine Signifikanz, womit die Entwicklungen statistisch zufällig bleiben.

Ergebnisse

Tabelle 50. Wachstumseffekte (Modell D) in der Zusammenarbeit in der Fachkonferenz und der Kooperation im Team, Studie 2, Lehrkräfte

		Parameter	Zusammenarbeit	Kooperation
<i>Feste Effekte</i>				
Ausgangswert	Konstante	γ_{00}	2.517*** (.147)	3.146*** (.142)
π_{0i}	Gruppe	γ_{01}	.310~ (.160)	.321* (.159)
	Bundesland (0=VS, 1=PS)	γ_{02}	-.313* (.211)	-.499** (.158)
Wachstumswert	Konstante	γ_{10}	.004 (.019)	-.018 (.015)
π_{1i}	Gruppe	γ_{11}	-.025 (.019)	-.004 (.017)
	Bundesland (0=VS, 1=PS)	γ_{12}	.021 (.024)	.025 (.018)
<i>Zufallseffekte</i>				
	innerhalb Personen	σ^2_{ε}	.211	.135
Level 1	Ausgangswert	σ^2_0	.171***	.240***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, BE/BB = Berlin/Brandenburg, HE = Hessen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparameters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

6.2.2.3 Der kompetenzorientierte Mathematikunterricht

Für die Merkmale der Unterrichtsgestaltung bleibt zu beachten, dass die Auskünfte der Lehrkräfte in BE/BB aus Studie 1 sich im 3. MZP auf die Schülerschaft der Jahrgangsstufe 10 bezogen. Studie 2 beinhaltet zwar die gleichen Lehrkräfte, die Unterrichtsgestaltung bezieht sich nunmehr jedoch auf eine andere Schülerklientel der neunten Jahrgangsstufe. Wenngleich die Initiierung der verschiedenen kompetenzorientierten Tätigkeiten nicht in Gänze losgelöst von den inhaltlichen Schwerpunktsetzungen zu betrachten ist, zeigt sich doch, dass die Lehrkräfte das in Studie 1 entwickelte Niveau im mathematischen Argumentieren auch in den neuen Klassen halten (Abschnitt 6.2.1.3 und Tabelle 51).

Obwohl grafisch deutliche Veränderungen erkennbar sind, können diese statistisch nicht nachgewiesen werden. Somit bleiben die Veränderungen zufällig, eine Varianzanalyse mit Messwiederholung ergab keine signifikanten Effekte. Es bleibt jedoch auffällig, dass die Schulen in BE/BB am Schuljahresanfang diese Kompetenz deutlich weniger einfordern als die Schulen in Hessen.

Ergebnisse

Tabelle 51. Deskriptive Statistiken zum Merkmal mathematisch Argumentieren

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>mathematisch Argumentieren (global)</i>								
Projektgruppe	46	4.32 (0.12)			46	4.33 (0.09)		
Vergleichsgruppe	38	4.11 (0.13)	1.37	.01	38	4.32 (0.12)	.01	.00
gesamt	84	4.23 (0.09)			84	4.32 (0.07)		
<i>mathematisch Argumentieren (BB/BB)</i>								
Projektgruppe	12	4.04 (0.28)			12	4.35 (0.21)		
Vergleichsgruppe	17	4.00 (0.16)			17	4.13 (0.11)		
<i>mathematisch Argumentieren (Hessen)</i>								
Projektgruppe	34	4.42 (0.14)	1.30	.05	34	4.32 (0.10)	0.79	.03
Vergleichsgruppe	21	4.21 (0.19)			21	4.47 (0.19)		
gesamt	84	4.23 (0.09)			84	4.32 (0.07)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), *F*-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Betrachtet man nur die Projektschulen, dann ist eine gegenläufige Veränderung zu beobachten. Die Projektschulen in Hessen fordern das mathematische Argumentieren zum zweiten Messzeitpunkt weniger ein, während die Projektschulen in BE/BB diesen Kompetenzbereich weiter ausbauen. Gründe hierfür könnten in dem IQB-Implementationskonzept liegen, das einen Schwerpunkt auf den kompetenzorientierten Unterricht legt. Dem steht entgegen, dass die hessischen Vergleichsschulen eine ähnliche Veränderung abbilden, ohne einen externen Input mit dem Schwerpunkt einer kompetenzorientierten Unterrichtsgestaltung.

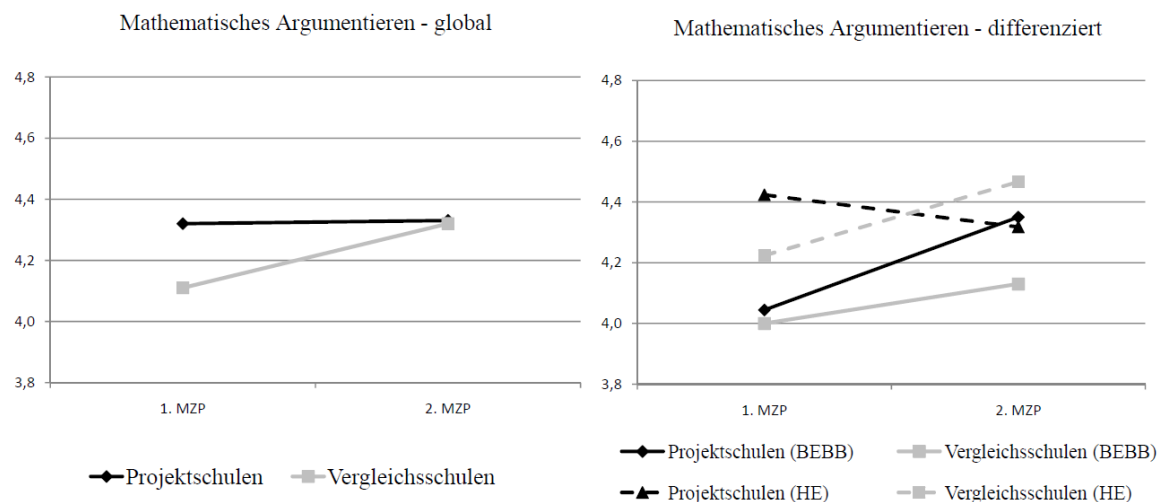


Abbildung 40. Mathematisch Argumentieren - Veränderung; Gegenüberstellung global und differenziert.

Die in den Bildungsstandards formulierte mathematische Kompetenz des Probleme lösens zeigt sich in der globalen Betrachtung stabil über das Schuljahr mit einer gleichbleibenden Steigerung in beiden Gruppen. Sowohl die Projekt- als auch die Vergleichsschulen fordern im Laufe eines Schuljahres in gleichem Maße häufiger solche mathematischen Tätigkeiten, die es erfordern Probleme zu lösen. Die Gruppenunterscheidung im ersten Messzeitpunkt ist statistisch unbedeutend, so auch die Veränderung der Gruppen über die Zeit. Es zeigt sich lediglich ein schwacher Haupteffekt der Gruppe, wodurch sehr vorsichtig von einer Verbesserung der Projektschulen gegenüber den Vergleichsschulen gesprochen werden kann ($F(1, 82) = 3.24, p = .076$).

Tabelle 52. Deskriptive Statistiken zum Merkmal *Probleme mathematisch Lösen*

Merkmal	MZIP 1				MZIP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>Probleme mathematisch Lösen (global)</i>								
Projektgruppe	46	4.48 (0.10)			46	4.60 (0.09)		
Vergleichsgruppe	38	4.23 (0.13)	2.53	.03	38	4.36 (0.10)	2.84	.03
gesamt	84	4.37 (0.08)			84	4.49 (0.07)		
<i>Probleme mathematisch Lösen (BB/BB)</i>								
Projektgruppe	12	4.29 (0.22)			12	4.63 (0.23)		
Vergleichsgruppe	17	4.33 (0.17)			17	4.26 (0.11)		
<i>Probleme mathematisch Lösen (Hessen)</i>								
Projektgruppe	34	4.55 (0.12)	1.43	.05	34	4.59 (0.10)	1.17	.04
Vergleichsgruppe	21	4.15 (0.19)			21	4.44 (0.18)		
gesamt	84	4.37 (0.08)			84	4.49 (0.07)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In der differenzierten Betrachtung auffallend, dass die Schulen in BE/BB eine annähernd einheitliche Unterrichtsgestaltung in Bezug auf das Lösen mathematischer Probleme teilen, während die Schulen in Hessen auffallend unterschiedliche Angaben zur Initiierung dieser Kompetenz angeben. Der Unterschied in den Anfangswerten der hessischen Schulen kann statistisch auch nachgewiesen werden ($F(1,54)=3.88, p = .054$). Hinsichtlich dieses Merkmals liegt für das Bundesland Hessen demnach eine Negativauswahl an Schulen vor. Die zum ersten Messzeitpunkt bereits hoch eingestuften Projektschulen in Hessen können dann auch im Laufe des Schuljahres diese Kompetenz nicht weiter ausbauen. Die hessischen Vergleichsschulen hingegen bauen diese Kompetenz im Verlauf eines Schul-

jahres weiter aus. Die Lehrkräfte in den Projektschulen in BE/BB bauen diese Kompetenz im Verlauf eines Schuljahres stark aus. Gründe könnten in den Inhalten des IQB-Implementationskonzeptes liegen. Die Lehrkräfte in den Vergleichsschulen verharren auf dem Anfangsniveau. Die Entwicklung der beiden Gruppen in BE/BB wird durch den Interaktionseffekt einer Varianzanalyse mit Messwiederholung auf einem Faktor qualifiziert ($F(1,27)=3.99, p=.056$). Die Projektgruppe initiiert damit systematisch häufiger die mathematische Kompetenz des Probleme Lösen.

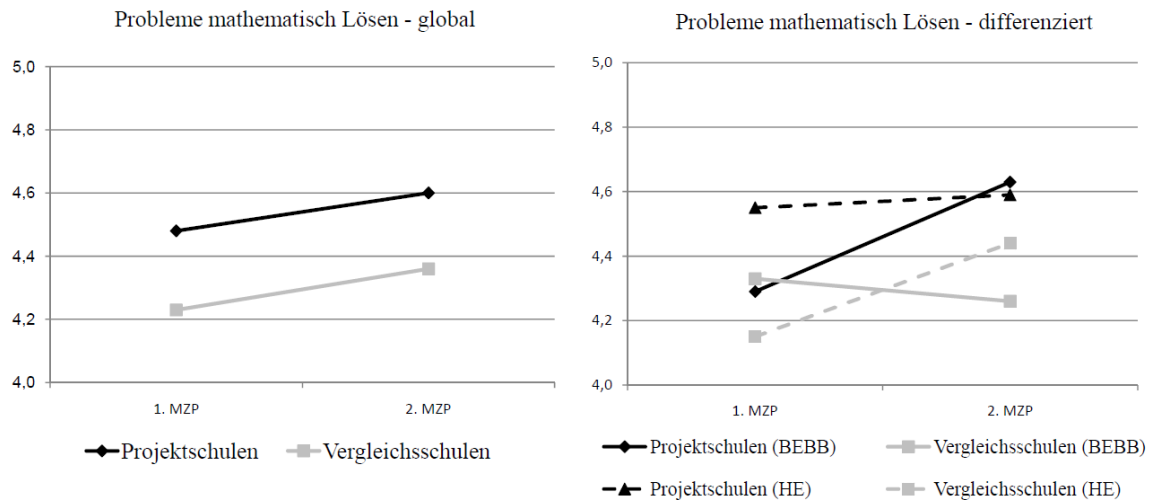


Abbildung 41. Probleme mathematisch Lösen - Veränderung; Gegenüberstellung global und differenziert.

Auch in Bezug auf die Tätigkeiten des mathematischen Modellierens geben die Lehrkräfte in den Projektschulen an, diese häufiger zu initiieren als die Lehrkräfte in den Vergleichsschulen. Allgemein ist besonders grafisch erkennbar, dass sich beide Gruppen im Verlauf eines Schuljahres dieser Kompetenz verstärkt annehmen und diese häufiger von den Schülerinnen und Schülern einfordern. Die Veränderung der beiden Gruppen kann einer Varianzanalyse mit Messwiederholung jedoch nicht standhalten und zeigt keine Ergebnisse. Auch der Unterschied zwischen den beiden Gruppen ist statistisch nicht bedeutsam. In der differenzierten Darstellung und mit Blick auf die Evaluation des IQB-Konzeptes zeigt sich zunächst, dass die hessischen Schulen die Tätigkeiten des mathematischen ganz unterschiedlich bewerten. Die Lehrkräfte in den hessischen Vergleichsschulen initiieren nach eigener Auskunft sehr viel weniger häufig diese mathematischen Tätigkeiten als ihre Kolleginnen und Kollegen der Projektschulen. Der Unterschied zwischen den hessischen Gruppen wird durch einen signifikanten Mittelwertunterschied bestätigt ($F(1, 54)= 4.40, p=.041$).

Ergebnisse

Tabelle 53. Deskriptive Statistiken zum Merkmal mathematisch Modellieren

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>mathematisch Modellieren (global)</i>								
Projektgruppe	46	4.23 (0.12)			46	4.35 (0.10)		
Vergleichsgruppe	38	3.97 (0.14)	2.00	.02	38	4.11 (0.14)	2.10	.03
gesamt	84	4.11 (0.09)			84	4.24 (0.08)		
<i>mathematisch Modellieren (BB/BB)</i>								
Projektgruppe	12	3.92 (0.24)			12	4.27 (0.20)		
Vergleichsgruppe	17	4.14 (0.17)			17	4.20 (0.16)		
<i>mathematisch Modellieren (Hessen)</i>								
Projektgruppe	34	4.33 (0.13)	1.82	.06	34	4.37 (0.11)	0.89	.03
Vergleichsgruppe	21	3.83 (0.22)			21	4.04 (0.21)		
gesamt	84	4.11 (0.09)			84	4.24 (0.08)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In den Schulen der Bundesländer BE/BB ist von einer einheitlichen Unterrichtsgestaltung im Hinblick auf die mathematische Kompetenz des Modellierens auszugehen. Die in Abbildung 42 sichtbaren Veränderungen in der differenzierten Darstellung können statistisch jedoch nicht bestätigt werden. Eine Varianzanalyse mit Messwiederholung auf einem Faktor über alle vier Gruppen brachte keine signifikanten Ergebnisse. Dennoch kann mit den rein deskriptiven Ergebnissen gezeigt werden, dass die Lehrkräfte in den Projektschulen BE/BB diese Kompetenz im Vergleich zu den Vergleichsschulen BE/BB und den Projektschulen HE weiter ausbauen konnten.

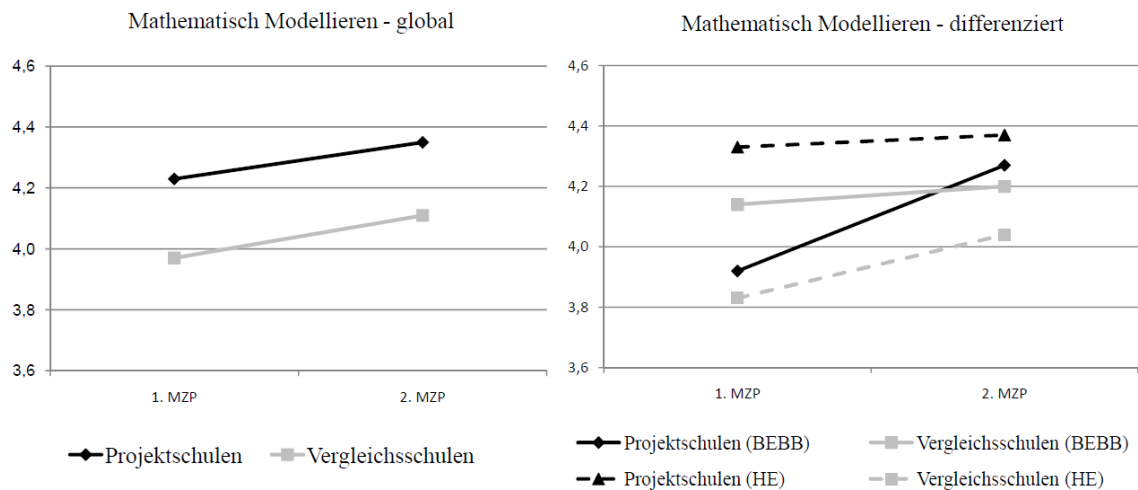


Abbildung 42. Mathematisch Modellieren - Veränderung; Gegenüberstellung global und differenziert.

Hier könnte ein Effekt des IQB-Implementationskonzeptes vorliegen. Dagegen spricht allerdings, dass die Vergleichsschulen in Hessen eine ähnliche Entwicklung zeigen, wenngleich die Veränderung schwächer ausfällt.

Im Merkmal „Darstellungen Verwenden“ ist in der Globaldarstellung keine auffallende Veränderung sichtbar, beide Gruppen, d.h. alle befragten Lehrkräfte verwenden im Laufe eines Schuljahres im Mittel häufiger mathematische Darstellungen in Form von Diagrammen und Tabellen. Die Gruppen zeigen dabei keinen Unterschied. Das Ausmaß der Veränderung ist in beiden Gruppen gleich groß. Wie grafisch bereits ersichtlich, kann eine systematisch unterschiedliche Entwicklung statistisch nicht nachgewiesen werden. Unterschiede sind jedoch in der differenzierten Darstellung erkennbar, wenngleich die Effekte der Gruppenunterschiede klein bleiben (Tabelle 54). Mittelwertunterschiede der einzelnen Gruppen zum ersten Messzeitpunkt können nicht nachgewiesen werden.

Tabelle 54. Deskriptive Statistiken zum Merkmal Darstellungen Verwenden

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>mathematische Darstellungen Verwenden (global)</i>								
Projektgruppe	46	3.91 (0.11)			46	4.00 (0.10)		
Vergleichsgruppe	38	3.85 (0.14)	0.12	.00	38	3.94 (0.12)	0.15	.00
gesamt	84	3.88 (0.09)			84	3.97 (0.08)		
<i>mathematische Darstellungen Verwenden (BB/BB)</i>								
Projektgruppe	12	3.76 (0.18)			12	4.05 (0.20)		
Vergleichsgruppe	17	4.10 (0.18)			17	4.15 (0.16)		
<i>mathematische Darstellungen Verwenden (Hessen)</i>								
Projektgruppe	34	3.96 (0.14)	1.25	.05	34	3.98 (0.11)	1.00	.04
Vergleichsgruppe	21	3.64 (0.20)			21	3.77 (0.18)		
gesamt	84	3.88 (0.09)			84	3.97 (0.08)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Grafisch zeigt sich jedoch eine Entwicklung der Projektschulen BE/BB, während die Projektschulen in Hessen auf ihrem Niveau verharren. Die Vergleichsschulen steigern die kompetenzorientierten Tätigkeiten des Darstellungen verwenden um wenige Skaleneinheiten. Am Ende des Schuljahres ist diese Kompetenz in den Bundesländern BE/BB höher ausgeprägt als im Bundesland Hessen.

Ergebnisse

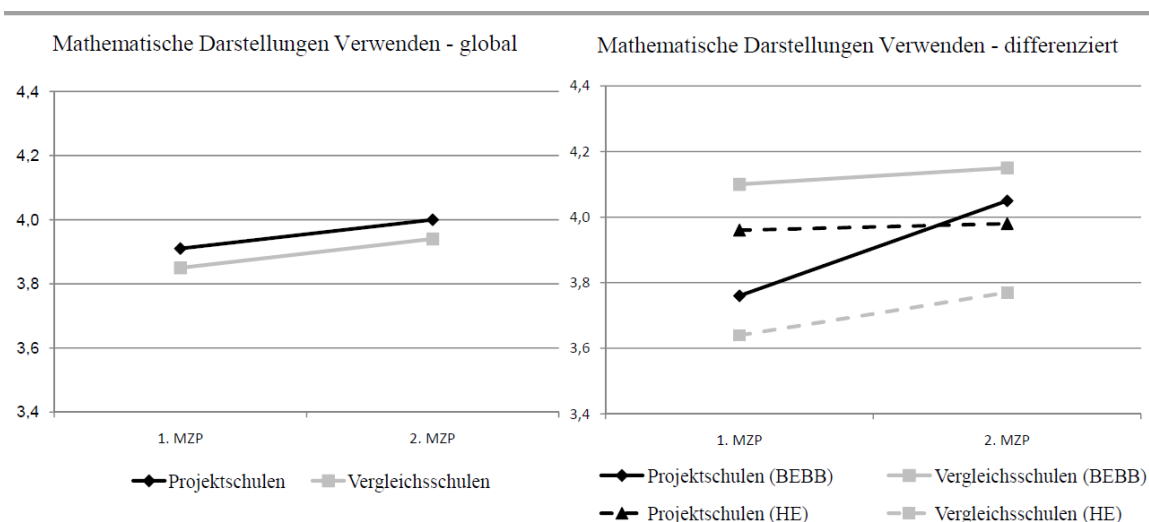


Abbildung 43. Mathematische Darstellungen Verwenden - Veränderung; Gegenüberstellung global und differenziert.

Das den deutschen Mathematikunterricht stark kennzeichnende Merkmal des technischen Arbeitens wird auch von den teilnehmenden Schulen als dominierend rückgemeldet. Alle Schulen befinden sich auf einem hohen Niveau von mehr als 4.5 Skalenpunkten. Im Verlauf eines Schuljahres zeigt sich, dass die Projektschulen darauf achten, diese mathematische Kompetenz nicht weiter zu fördern, während die Vergleichsschulen im Allgemeinen das routinemäßige Abarbeiten von mathematischen Formeln weiter ausbauen und verstärkt initiieren.

Tabelle 55. Deskriptive Statistiken zum Merkmal "technisch Arbeiten"

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>technisch Arbeiten (global)</i>								
Projektgruppe	46	4.72 (0.09)			46	4.55 (0.09)		
Vergleichsgruppe	38	4.54 (0.10)	2.11	.03	38	4.62 (0.09)	0.35	.00
gesamt	84	4.64 (0.07)			84	4.58 (0.06)		
<i>technisch Arbeiten (BB/BB)</i>								
Projektgruppe	12	4.67 (0.19)			12	4.65 (0.11)		
Vergleichsgruppe	17	4.59 (0.15)			17	4.51 (0.15)		
<i>technisch Arbeiten (Hessen)</i>								
Projektgruppe	34	4.74 (0.10)	0.83	.03	34	4.51 (0.11)	0.71	.03
Vergleichsgruppe	21	4.49 (0.13)			21	4.71 (0.12)		
gesamt	84	4.64 (0.07)			84	4.58 (0.06)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In der globalen Darstellung der Projekt- und Vergleichsschulen führt das Sinken der Häufigkeit dieser Tätigkeiten bei den Projektschulen und der leichte Anstieg in den Vergleichsschulen zu einem schwachen Interaktionseffekt der Gruppe mit der Zeit ($F(1, 82) = 3.61, p < .10$). Vorsichtig interpretiert veränderten die Lehrkräfte in den Projektschulen ihre Unterrichtsgestaltung im Hinblick auf dieses Merkmal systematisch im Vergleich zu den Lehrkräften in den Vergleichsschulen.

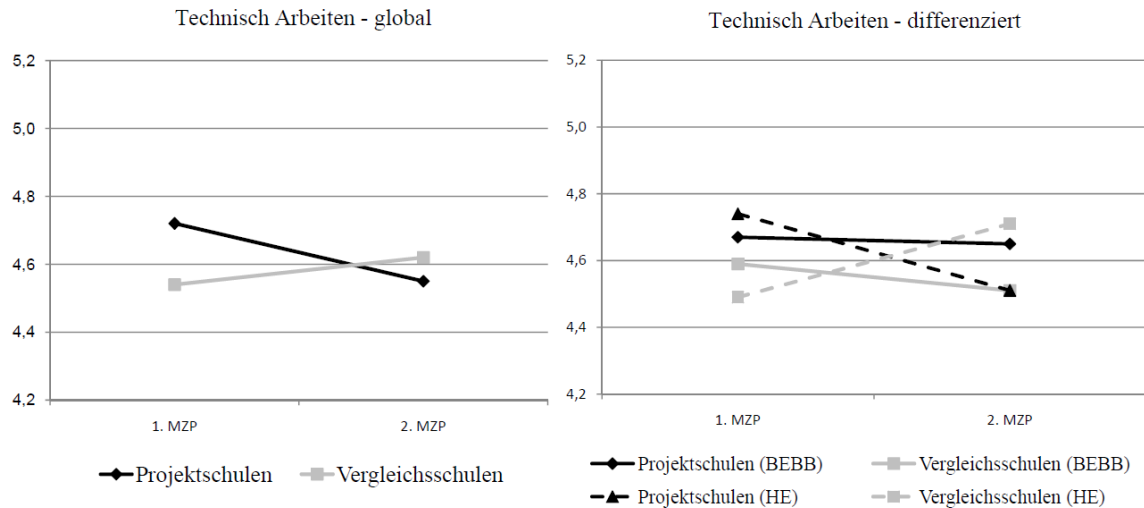


Abbildung 44. Technisch Arbeiten - Veränderung; Gegenüberstellung global und differenziert.

Werden die Daten weiter differenziert nach dem Bundesland und der Intervention, wird erkennbar, dass die globale Veränderung hauptsächlich durch die Unterrichtsveränderung in den hessischen Projektschulen herbeigeführt wurde. Die Projektschulen aus BE/BB verändern ihre Unterrichtsgestaltung nicht, die Vergleichsschulen BE/BB zeigen vielmehr die gewünschte Veränderung in Richtung Abbau der routinierten mathematischen Arbeitsweise. Die hessischen Vergleichsschulen bauten diese Kompetenz im Verlauf des Schuljahres weiter aus. In ihren Ausgangswerten unterscheiden sich die verschiedenen Gruppen nicht, der Interaktionseffekt aus einer Varianzanalyse mit Messwiederholung bleibt jedoch erhalten: $F(3, 80) = 3.61, p < .10$.

Das Merkmal des mathematischen Kommunizierens ist ebenfalls auf einem hohen Niveau im deutschen Mathematikunterricht verankert. Die Gruppen unterscheiden sich statistisch nicht in ihren Anfangswerten, zeigen jedoch auch keine Veränderung. Mit Blick auf die deskriptiven Statistiken ist erkennbar, dass die Projektschulen insgesamt (BE/BB und Hessen) ihre Schülerinnen und Schüler nicht weniger häufig auffordern, mathematisch zu kommunizieren. Die Vergleichsschulen steigern die Häufigkeit dieser Tätigkeiten im Verlauf eines Schuljahres, reichen jedoch nicht an das Niveau der Projektschulen heran.

Ergebnisse

Tabelle 56. Deskriptive Statistiken zum Merkmal mathematisch Kommunizieren

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>mathematisch Kommunizieren (global)</i>								
Projektgruppe	46	4,85 (0,10)			46	4,88 (0,09)		
Vergleichsgruppe	38	4,67 (0,12)	1,38	.02	38	4,78 (0,12)	0,48	.01
gesamt	84	4,77 (0,08)			84	4,83 (0,07)		
<i>mathematisch Kommunizieren (BB/BB)</i>								
Projektgruppe	12	4,69 (0,24)			12	5,08 (0,19)		
Vergleichsgruppe	17	4,66 (0,18)			17	4,64 (0,13)		
<i>mathematisch Kommunizieren (Hessen)</i>								
Projektgruppe	34	4,91 (0,10)	0,77	.03	34	4,81 (0,10)	1,07	.04
Vergleichsgruppe	21	4,69 (0,16)			21	4,89 (0,19)		
gesamt	84	4,77 (0,08)			84	4,83 (0,07)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

In der differenzierten Betrachtung wird deutlich, dass sich die Lehrkräfte in den Projekt-schulen BE/BB dieser mathematischen Kompetenz annehmen und sie um 0,4 Skalen-punkte weiter ausbauen, während die Vergleichsschulen BE/BB keine Veränderung zeig-en. Die hessischen Projekt-schulen konnten das anfänglich sehr hohe Niveau über das Schuljahr hinweg nicht halten und zeigen in der Initiierung von Schülertätigkeiten im mathematischen Kommunizieren eine leicht rückläufige Tendenz. Ihre Kolleginnen und Kollegen aus den Vergleichsschulen in Hessen steigern die Aktivitäten in diesem Bereich und zeigen am Ende des Schuljahres ein höheres Niveau als die Projekt-schulen in Hessen. Statistisch bleiben die grafisch sichtbaren Veränderungen jedoch zufällig.

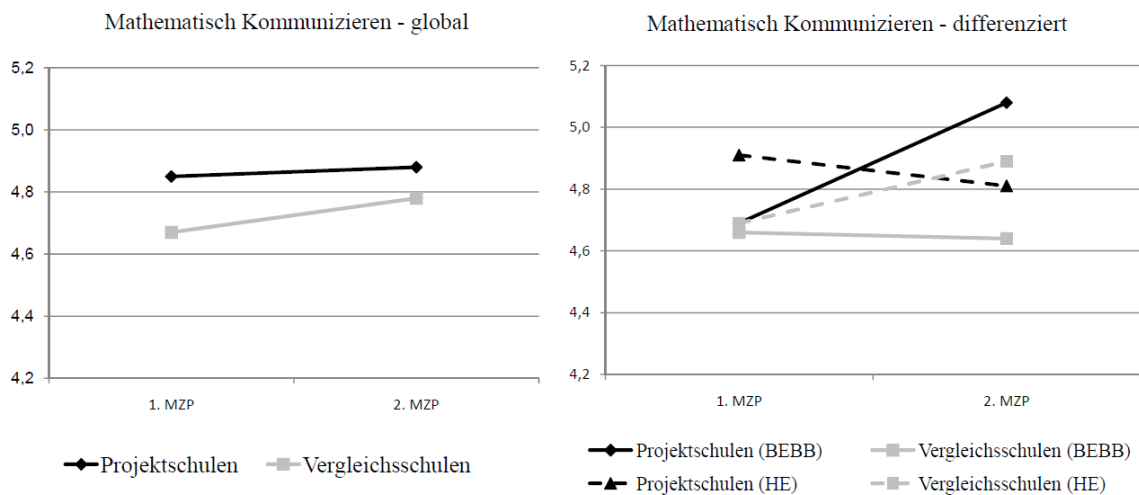


Abbildung 45. Mathematisch Kommunizieren - Veränderung; Gegenüberstellung global und differenziert.

Werden die unterschiedlichen Erhebungszeitpunkte in Rechnung gestellt und soll eine Gruppenunterscheidung ermöglicht werden, dann stellen Mehrebenenmodelle das geeignete Mittel dar, um statistische Bedeutsamkeiten sichtbar zu machen. In Tabelle 57 zeigt sich, dass vier der sechs untersuchten Merkmale eines kompetenzorientierten Unterrichts keine statistisch signifikanten Auffälligkeiten aufweisen. Die Unterrichtsgestaltung in den Bereichen des mathematischen Probleme lösens, Modellierens, Darstellungen verwenden und Kommunizierens variiert, jedoch so zufällig, dass bedeutende Unterschiede nicht analysiert werden können. Somit ist im Verlauf eines Schuljahres für diese Merkmale auch keine statistisch signifikante Entwicklung zu beobachten. Mit Blick auf die Konstanten als mittlerer Skalenwert einer Durchschnittslehrkraft in der Stichprobe rangieren die Skalenwerte in ihrer Bedeutung um den Wert vier, der mit der nachgefragten Häufigkeit von „jeden Monat“ assoziiert wird⁵⁹. Einzig die Gestaltung im mathematischen Kommunizieren weist einen signifikanten Wachstumswert der Vergleichsschulen auf. Bereits zum ersten Messzeitpunkt zeigen die Schulen aus Hessen einen schwach signifikant höheren Anfangswert auf als die Schulen aus Berlin/Brandenburg. Die interessierende Veränderung gelingt nur bedeutend für die Vergleichsschulen in Berlin/Brandenburg. Das Wachstum der Projektschulen hingegen unterscheidet sich in ihrer Veränderung gegenüber dem Anfangswert nicht bedeutend.

Im Merkmal des technischen Arbeitens stellt sich eine ungleich Ausgangslage dar. Die Lehrkräfte in den Projektschulen in Berlin/Brandenburg initiieren Tätigkeiten des routinierten Abarbeitens signifikant häufiger als die Lehrkräfte in den Vergleichsschulen. Das Ziel sollte daher ein Abbau des technischen Arbeitens im Mathematikunterricht sein. Mit Blick auf die Veränderung zum zweiten Messzeitpunkt haben die Lehrkräfte der Projektschulen es auch verstanden, diese zwar notwendigen, aber bereits dominierenden Tätigkeiten im Unterricht signifikant zu verringern (Tabelle 57).

⁵⁹ Skalierung beim Fragebogen zu den kompetenzorientierten Tätigkeiten: 1=nie, 2= 1 Mal pro Schuljahr, 3 = mehrere Male pro Schuljahr, 4 = jeden Monat, 5 = jede Woche, 6 = fast jede Stunde

Ergebnisse

Tabelle 57. Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 2, Lehrkräfte

		Parameter	Argumentieren	Probleme Lösen	Modellieren	Darstellungen	technisch Arbeiten	Kommunizieren
<i>Feste Effekte</i>								
Ausgangswert	Konstante	γ_{00}	3.748*** (.203)	4.104*** (.209)	3.806*** (.230)	3.797*** (.226)	4.519*** (.180)	4.485*** (.226)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	.301 (.224)	.331 (.219)	.301 (.247)	.164 (.235)	.381* (.171)	.217 (.218)
	Bundesland (0= BE/BB, 1=HE)	γ_{02}	.420~ (.242)	.071 (.228)	.140 (.270)	-.051 (.259)	-.072 (.194)	.208 (.245)
Wachstumswert	Konstante	γ_{10}	.049* (.019)	.023 (.021)	.032 (.025)	.027 (.024)	.008 (.020)	.030 (.025)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	-.034 (.021)	-.014 (.022)	-.008 (.025)	-.008 (.023)	-.046* (.019)	-.014 (.026)
	Bundesland (0= BE/BB, 1=HE)	γ_{12}	-.028 (.028)	-.003 (.022)	-.019 (.028)	-.017 (.028)	.011 (.021)	-.022 (.027)
<i>Zufallseffekte</i>								
	innerhalb Personen	σ^2_{ε}	.251	.259	.324	.293	.193	.328
Level 1	Ausgangswert	σ^2_0	.308***	.232***	.317***	.284***	.151***	.151***

Anmerkungen. VS= Vergleichsschulen, PS=Projektschulen, HE= Hessen, BE/BB= Berlin/Brandenburg; imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparameters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

6.2.2.4 Zusammenfassung der Ergebnisse aus Studie 2

Die Ergebnisse der Studie lassen sich folgendermaßen zusammenfassen:

- Die Grundeinstellungen gegenüber den Bildungsstandards konnten in beiden Bundesländern (räumlich) und durch beide Unterrichtsentwicklungskonzepte (inhaltlich) nicht verändert werden.
- Die Intensität der Zusammenarbeit in der Fachkonferenz sowohl inhaltlich (Kooperation) als auch organisatorisch (Zusammenarbeit in der Fachkonferenz) können nicht weiter intensiviert werden.
- Die Projektarbeit im Bundesland Hessen führt zu einem kaum veränderten Unterrichtsangebot, deren Schwerpunkt auf den Kompetenzen des mathematischen Probleme lösens, modellierens und mathematischen kommunizierens liegt.
- Die Projektarbeit in den Bundesländern BE/BB führt zu einer starken Angebotszunahme in allen mathematischen Kompetenzbereichen. Das vorherrschende technische Arbeiten wurde nicht weiter ausgebaut. Trägt man der Ausgangssituation der Projektschulen in BE/BB Rechnung, veränderte sich das Unterrichtsangebot dahingehend, dass häufiger kompetenzorientierte Tätigkeiten in den Bereichen des mathematischen Argumentierens, Probleme lösens, modellierens, Darstellungen verwenden und kommunizierens eingefordert werden.

6.2.3 Zusammenfassung der Forschungsfrage 1

Ziel des Projektes war es, die Gestaltung des Mathematikunterrichts auf eine verstärkte Kompetenzorientierung hin zu verändern. Bei den Lehrkräften sollten Veränderungsprozesse angestoßen werden, die zu einem Mathematikunterricht führen, der gemeinsam vom Kollegium (Fachkonferenz) getragen wird und auf die Ausbildung bzw. den Aufbau der in den Bildungsstandards beschriebenen Kompetenzen fokussiert. Bei erfolgreicher Arbeit sollten sich die Einstellungen gegenüber den Bildungsstandards positiv entwickeln bzw. die Ängste abgebaut werden. Die Ergebnisse der Arbeit in den Fachkonferenzen zeigen, dass die Lehrkräfte in den Projektschulen in einigen Merkmalen Veränderungen zeigten. Zusammenfassend lassen sich folgende Befunde festhalten:

- Lehrkräfte, die einer Bildungsinnovation kritisch gegenüberstehen und weniger stark im Team kooperieren, lehnen Hilfe bei der Implementation ab.
- Der Unterricht der teilnehmenden Lehrkräfte ist durch das Abarbeiten von mathematischen Routinen geprägt.

- Die Arbeit in den Fachkonferenzen der Projektschulen führte zwangsläufig zu häufigeren Treffen der Lehrkräfte. Als Folge verbesserte sich die Kooperation in den Fachkonferenzen der Projektschulen systematisch.
- Der Unterricht in den Projektschulen erfuhr eine stärkere Orientierung hin zum mathematischen argumentieren, modellieren und Probleme lösen.
- Die Gegenüberstellung von Querschnittsdaten mit den Längsschnittdaten der Studie 1 zeigt, dass eine Verbreitung des neuen Wissens aus der Arbeit im Projekt nicht stattfand. Die gewünschten Synergieeffekte der Lehrkräfte des Längsschnitts mit denen des Querschnitts konnten nicht beobachtet werden. Einzig den Lehrkräften, die stetig im Projekt arbeiteten, gelang es, den Input und die Möglichkeiten aus dem Projekt in den Unterricht zu tragen und häufiger Tätigkeiten zu initiieren, die dem Kompetenzgedanken eines modernen Mathematikunterrichts Rechnung tragen.
- Das Implementationskonzept des IQB wird für die Zielgruppe der Lehrkräfte positiv evaluiert. Gegenüber den Hessischen Kolleginnen und Kollegen, die am Projekt KOU teilnahmen, zeigen sie in allen Bereichen mit Ausnahme des technischen Arbeitens eine positive Entwicklung.

6.3 Forschungsfrage 2: Leistungs- und Wahrnehmungsveränderungen von Schülerinnen und Schülern aus einer kompetenzorientierten Perspektive

Neben der Gestaltung eines kompetenzorientierten Mathematikunterrichts sind es am Ende die Schülerinnen und Schüler, die ihre Kompetenzen nach einem vorherigen Angebot aufbauen und anwenden sollen. Mit Blick auf die Anstrengungen der Bundesländer im Bereich der Implementation von schulischen Reformen bleibt am Ende die Frage, was beim Schüler angekommen ist, wenn die nötigen Voraussetzungen geschaffen wurden. Für den Erwerb und den Aufbau mathematischer Kompetenzen ist ein kompetenzorientierter Unterricht als Angebot der Lehrkräfte Voraussetzung. Wie sich die Veränderung im Unterrichtsangebot auf Seiten der Lehrkräfte bei den Schülerinnen und Schülern niederschlägt, soll im Folgenden untersucht werden.

6.3.1 Studie 1 „Erprobung eines Implementationskonzeptes in Berliner und Brandenburger Schulen“:

Die zeitgleich mit den Lehrkräften erhobenen Daten der Schülerinnen und Schüler bilden ab, inwieweit vorgenommene Änderungen in der Unterrichtsgestaltung bei den Empfängern ankommen. Im Idealfall unterscheiden sich die Schülergruppen zum Beginn der Befragung nicht voneinander und zeigen im Verlauf eine Unterschiedlichkeit, die auf veränderte Wahrnehmungen der Schülerinnen und Schüler in den Projektschulen zurückzuführen ist. Diese veränderten Wahrnehmungen könnten über die Intentionen der Lehrkräfte auf das Wirken des Implementationskonzeptes zurückgeführt werden.

6.3.1.1 Entwicklung des mathematischen Selbstkonzeptes

Das auf die mathematische Leistung wirkende Selbstkonzept zeigt im Querschnitt anfänglich keine Unterschiede zwischen den Schülerinnen und Schülern der Projekt- und Vergleichsschulen. Zu beobachten ist jedoch eine signifikante Gruppenunterscheidung im 3. Messzeitpunkt (Querschnitt), die sich jedoch aufhebt, wenn man nur die Schülerinnen und Schülern betrachtet, die zu allen drei Datenerhebungen anwesend waren. Mit Blick auf die Mittelwerte der Skala des mathematischen Selbstkonzeptes konnten die Schülerinnen und Schüler in den Vergleichsschulen ihr Selbstkonzept steigern, während die Lernerinnen und Lerner in den Projektschulen auf einem Niveau unter dem theoretischen Mittelwert der Skala verblieben. Tendenziell stehen diese Schülerinnen und Schüler dem Fach Mathematik eher kritisch gegenüber. Die Vergleichsschulen erreichen einen nur wenig höheren Mittelwert, der sich ebenfalls um den theoretischen Mittelwert der Skala positioniert. Eine Unterscheidung der Stichproben des Quer- und Längsschnittes zeigt auch, dass in beiden Gruppen diejenigen Schülerinnen und Schüler an der Studie teilnahmen, die ein höheres Selbstkonzept aufweisen als ihre Mitschüler mit fehlenden Datenerhebungen. Dies kommt in den höheren Mittelwerten des Längsschnitts zum Ausdruck (Tabelle 58). Die Befunde werden durch die grafische Darstellung noch einmal verdeutlicht (Anhang E). Im Vergleich von Quer- und Längsschnitt sind die Schülerinnen und Schüler des Längsschnitts sowohl der Vergleichs- als auch der Projektschulen auf einem höheren Niveau des mathematischen Selbstkonzeptes zu verorten. Die Entwicklung weist jedoch schwach tendenziell darauf hin, dass die Schülerinnen und Schüler der Projektschulen in ihrem Gefühl dem Fach Mathematik gegenüber unsicherer werden. Diese Unsicherheit könnte auf die Verwendung neuer Aufgabentypen im Mathematikunterricht zurückzuführen sein.

Ergebnisse

Tabelle 58: Deskriptive Statistiken zum Merkmal mathematisches Selbstkonzept, Schülerstichprobe

	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
mathematisches Selbstkonzept (Querschnitt)															
Projektgruppe	760	2.42 (0.76)				636	2.43 (0.76)				736	2.40 (0.78)			
Vergleichsgruppe	585	2.47 (0.73)		1.21	.00	460	2.51 (0.74)		2.57	.00	563	2.49 (0.75)		5.11*	.00
gesamt	1345	2.44 (0.75)				1096	2.47 (0.75)				1299	2.44 (0.77)			
mathematisches Selbstkonzept (Längsschnitt)															
Projektgruppe	487	2.50 (0.74)				481	2.48 (0.75)				481	2.44 (0.77)			
Vergleichsgruppe	297	2.52 (0.72)		0.21	.00	298	2.51 (0.74)		0.24	.00	296	2.50 (0.74)		1.23	.00
gesamt	784	2.51 (0.73)				779	2.49 (0.75)				777	2.46 (0.76)			
mathematisches Selbstkonzept (imputiert)															
Projektgruppe	720	2.46 (0.71)	.03			720	2.43 (0.74)	.03			720	2.39 (0.76)	.03		
Vergleichsgruppe	528	2.51 (0.66)	.03	1.89	.00	528	2.51 (0.71)	.03	3.04	.00	528	2.51 (0.72)	.03	6.48	.01
gesamt	1248	2.48 (0.69)	.02			1248	2.46 (0.73)	.02			1248	2.44 (0.74)	.02		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

6.3.1.2 Entwicklung des mathematischen Interesses

Etwas anders verhält es sich, wenn es um das mathematische Interesse geht. Zunächst weisen die Gruppenmittelwerte darauf hin, dass bei allen Schülerinnen und Schülern nur ein begrenztes Interesse an Mathematik entwickelt werden konnte. Alle Gruppenmittelwerte liegen unter dem theoretischen Mittelwert von 2.5 Skalenpunkten. Es zeigt sich weiterhin, dass die Schülerinnen und Schüler der Vergleichsschulen am Anfang ein signifikant größeres Interesse an Mathematik aufweisen als die Schülerinnen und Schüler in den Projektschulen. Betrachtet man jedoch nur die Personen, die an allen drei Erhebungen teilnahmen (Längsschnitt), dann verliert sich dieser Unterschied wieder. Folglich nahmen an der ersten Datenerhebung in den Vergleichsschulen besonders Schülerinnen und Schüler teil, die ein Interesse an Mathematik aufweisen. Im Verlauf der Studie können die Projekt- und Vergleichsschulen das Interesse an Mathematik bei den Schülerinnen und Schülern nicht steigern. Die Schülerinnen und Schüler in den Projektschulen berichten ein stetig fallendes Interesse am Mathematikunterricht vom ersten bis zum dritten Messzeitpunkt. Diese Entwicklung könnte mit dem veränderten Unterricht der Lehrkräfte zusammenhängen. Allerdings sinkt das Interesse an Mathematik auch in den Vergleichsschulen, zunächst etwas stärker und steigt wieder zum dritten Messzeitpunkt, jedoch ohne das Anfangsniveau zu erreichen. Aufgrund dieser Beobachtungen wird angenommen, dass das sinkende Interesse bei allen Schülerinnen und Schülern mit den angebotenen mathematischen Inhalten zu tun hat und weniger mit der Art und Weise, wie unterrichtet wird. Die anfänglich höheren Werte bei allen Schülerinnen und Schülern könnten auch durch eine sozial erwünschte Beantwortung der Items entstanden sein.

Ergebnisse

Tabelle 59. Deskriptive Statistiken zum Merkmal mathematisches Interesse, Schülerstichprobe

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
mathematisches Interesse (Querschnitt)															
Projektgruppe	760	2.11 (0.72)				636	2.05 (0.68)				735	2.03 (0.74)			
Vergleichsgruppe	585	2.21 (0.71)		6.67*	.01	460	2.08 (0.71)		0.36	.00	563	2.12 (0.70)		5.31*	.00
gesamt	1345	2.15 (0.72)				1096	2.06 (0.70)				1299	2.07 (0.73)			
mathematisches Interesse (Längsschnitt)															
Projektgruppe	487	2.13 (0.72)				481	2.09 (0.69)				481	2.05 (0.73)			
Vergleichsgruppe	297	2.16 (0.67)		0.49	.00	298	2.04 (0.69)		0.31	.00	296	2.08 (0.67)		0.47	.00
gesamt	784	2.14 (0.71)				779	2.07 (0.69)				777	2.06 (0.71)			
mathematisches Interesse (imputiert)															
Projektgruppe	719	2.15 (0.69)	.03			719	2.07 (0.66)	.03			719	2.05 (0.71)	.03		
Vergleichsgruppe	528	2.23 (0.65)	.03	5.36	.00	528	2.11 (0.69)	.03	0.99	.00	528	2.13 (0.69)	.03	3.94	.00
gesamt	1247	2.19 (0.67)	.02			1247	2.09 (0.67)	.02			1247	2.08 (0.70)	.02		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

6.3.1.3 Entwicklung der Lehrerwahrnehmung

Im Merkmal der Lehrerwahrnehmung ging es um die Bewertung der Mathematik Lehrkraft durch die Schülerinnen und Schüler. Die Items fragten weniger nach der Fachkenntnis des Lehrers, vielmehr sollte eine Einschätzung des Lehrers in seinem unterrichtlichen Auftreten abgegeben werden (z.B. Unser Mathematiklehrer unterrichtet mit Begeisterung.)

Anhand der Mittelwerte zur Skala der Lehrerwahrnehmung wird sichtbar, dass alle Schülerinnen und Schüler ihre Lehrkräfte schwach positiv bewerten. Die Mittelwerte liegen ausnahmslos über dem theoretischen Mittelwert der Skala von 2.5. Eine ANOVA zum ersten Messzeitpunkt ergab keine signifikanten Unterschiede in der Bewertung der Lehrkraft durch die Schülerinnen und Schüler (*Tabelle 60*). Im Verlauf der Studie reagieren die Schülerinnen und Schüler der Projektschulen jedoch auf das veränderte Verhalten der Lehrkraft. Die Zustimmung zur Skala der Lehrerwahrnehmung verringert sich in den Projektschulen um mehr als 0.1 Skalenpunkte. Daraus ergeben sich signifikante Unterschiede in den Gruppen zu den Messzeitpunkten zwei und drei, die jedoch entgegen den Erwartungen verlaufen. Die Schülerinnen und Schüler in den Projektschulen stufen die Sympathie für ihre Mathematik Lehrkraft herab. Im dritten Messzeitpunkt stagnieren die Werte auf diesem geringeren Niveau. Eine MANOVA bestätigt die Veränderung des Merkmals über die Zeit in einem sehr schwachen Haupteffekt ($F(2; 1510) = 2.34; p < .10$), ein bestätigender Interaktionseffekt zur Absicherung der unterschiedlichen Entwicklung in den beiden Gruppen wird jedoch nicht erreicht. Ein höheres Maß an Kompetenzorientierung könnte durch veränderte Frage- und Aufgabenstellungen dazu führen, dass sich die Schülerinnen und Schüler verunsichert fühlen und dieser Wahrnehmung durch die Lehrerbewertung Rechnung tragen.

Ein anderer Grund für diese unterschiedliche Entwicklung könnte in einer anfänglich von Seiten der Schülerinnen und Schüler produzierten sozialen Erwünschtheit liegen, die in den Vergleichsschulen aufrecht erhalten wurde.

Ergebnisse

Tabelle 60. Deskriptive Statistiken zum Merkmal Lehrerwahrnehmung, Schülerstichprobe

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
Lehrerwahrnehmung (Querschnitt)															
Projektgruppe	755	2.82 (0.79)				631	2.69 (0.86)				735	2.69 (0.88)			
Vergleichsgruppe	579	2.84 (0.83)		0.23	.00	453	2.83 (0.86)		6.96**	.01	562	2.90 (0.81)		20.47***	.02
gesamt	1334	2.83 (0.80)				1084	2.75 (0.86)				1297	2.78 (0.85)			
Lehrerwahrnehmung (Längsschnitt)															
Projektgruppe	484	2.87 (0.77)				477	2.69 (0.85)				482	2.72 (0.87)			
Vergleichsgruppe	296	2.89 (0.85)		0.04	.00	293	2.84 (0.85)		5.25*	.01	297	2.93 (0.82)		10.81**	.01
Gesamt	780	2.88 (0.80)				770	2.75 (0.85)				779	2.80 (0.86)			
Lehrerwahrnehmung (imputiert)															
Projektgruppe	719	2.81 (0.73)	.03			719	2.65 (0.82)	.03			719	2.68 (0.83)	.03		
Vergleichsgruppe	527	2.82 (0.74)	.03	0.10	.00	527	2.78 (0.81)	.04	8.26	.01	527	2.87 (0.77)	.03	17.25	.01
Gesamt	1246	2.81 (0.73)	.02			1246	2.71 (0.82)	.02			1246	2.76 (0.81)	.02		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Stellt man die unterschiedlichen Erhebungszeitpunkte mit Hilfe von Mehrebenenmodellen in Rechnung, dann bestätigen sich die bereits berichteten Befunde. Im mathematischen Selbstkonzept ist keine signifikante Entwicklung sowohl bei den Projekt- als auch bei den Vergleichsschulen zu berichten. Die Entwicklungen im Merkmal des mathematischen Interesses gehen systematisch in eine negative Richtung. Allerdings ist hier von unterschiedlichen Anfangswerten auszugehen. Die Schülerinnen und Schüler in den Projektschulen zeigen von Anfang an geringeres Interesse an Mathematik. In den Vergleichsschulen nimmt das Interesse signifikant ab, während in den Projektschulen keine bedeutende Entwicklung zu verzeichnen ist. In der Lehrerbewertung urteilen die Schülerinnen und Schüler der Projektschulen schlechter über ihre Mathematiklehrkraft als zu Beginn des Projektes. Die negative Entwicklung ist statistisch nachweisbar (Tabelle 61). Gründe dafür könnten auch in der auf die Schülerinnen und Schüler übertragene Überforderung in den Projektschulen liegen. Mit Blick auf die Itemformulierungen könnte eine Bewertung der Lehrkraft negativ ausfallen, wenn diese schlecht vorbereitet, überarbeitet und daher unkonzentriert oder auch unsicher in Bezug auf die zu erprobenden Aufgaben auftritt.

Tabelle 61. Wachstumseffekte (Modell D) zum mathematischen Selbstkonzept, Interesse an Mathematik und der Lehrerbewertung, Studie 1, Schülerinnen und Schüler

		Parameter	math. Selbstkonzept	Interesse	Lehrerbewertung
<i>Feste Effekte</i>					
Ausgangswert	Konstante	γ_{00}	2.518*** (.031)	2.224*** (.033)	2.792*** (.036)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	-.052 (.041)	-.080~ (.042)	-.001 (.046)
Wachstumswert	Konstante	γ_{10}	-.001 (.002)	-.009** (.003)	.004 (.003)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	-.004 (.002)	.002 (.003)	-.014** (.004)
<i>Zufallseffekte</i>					
	innerhalb Personen	σ^2_e	.109	.149	.253
Level 1	Ausgangswert	σ^2_0	.379***	.302***	.305***
Level 2	Wachstumswert	σ^2_1	.001***	.001***	.001***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

6.3.1.4 Entwicklung der mathematischen Kompetenzen

Inwiefern die von den Lehrkräften in den Projektschulen angegebene Steigerung der Häufigkeit bei der Initiierung von kompetenzorientierten Tätigkeiten auch von den Schülerinnen und Schülern wahrgenommen wird, sollen die folgenden Ergebnisse aus Schülersicht zeigen.

Hinsichtlich des mathematischen Argumentierens unterscheiden sich die Schülerinnen und Schüler in den Projekt- und Vergleichsschulen anfangs nicht statistisch voneinander. Im Querschnitt ergibt sich zum dritten Messzeitpunkt ein signifikanter Gruppenunterschied, der von den Teilnehmern des Längsschnittes jedoch nicht bestätigt wird. Ein Vergleich zwischen den Daten des Quer- und des Längsschnittes ergibt, dass in die Mittelwertberechnungen des Querschnitts in den Projektschulen Beurteilungen eingingen, die diese Tätigkeit im Unterricht nicht erkennen konnten und den Mittelwert damit nach unten verzerrten (Tabelle 62). Grundsätzlich bestätigen die Daten des Längsschnittes eine Negativauswahl der Projektschulen in diesem Merkmal, was durch den signifikanten Gruppenunterschied im ersten Messzeitpunkt des Längsschnittes bestätigt wird.

In der grafischen Darstellung (siehe Anhang E) ist gut erkennbar, dass im Mathematikunterricht der Vergleichsschulen häufiger Tätigkeiten, die dieser Kompetenz zuzuordnen sind, wahrgenommen werden. Diese Wahrnehmung ändert sich auch nicht, die Schülerinnen und Schüler in den Vergleichsschulen verbleiben über alle drei Messzeitpunkte auf einem höheren Niveau als ihre Mitschüler in den Projektschulen, anders als von Lehrkräften berichtet. Dennoch ist vor allem im Längsschnitt zu beobachten, dass die Schüler in den Projektschulen stetig häufiger Tätigkeiten des mathematischen Argumentierens wahrnehmen. Zum dritten Messzeitpunkt unterscheiden sich die beiden Gruppen nicht mehr signifikant voneinander, was durch die gesteigerten Mittelwerte in den Projektschulen begründet wird.

Im Vergleich des Quer- und Längsschnitts wird deutlich, dass in den Projektschulen diejenigen Schülerinnen und Schüler aus der Studie ausstiegen, die diese Kompetenz nur geringer wahrnahmen. Gleiches gilt für die Vergleichsschulen. Beide Mittelwerte des Querschnitts liegen unter den Werten des Längsschnitts, womit zum ersten Messzeitpunkt viele Schülerinnen und Schüler an der Studie teilnahmen, die das mathematische Argumentieren im Unterricht weniger häufig wahrnahmen. Die Betrachtung im Längsschnitt zeigt, dass die anfänglich signifikant unterschiedlichen Gruppen sich über die Zeit annähern, wobei die Schülerin und Schüler in den Projektschulen die von Lehrkräften häufiger

figer angebotenen Situationen zum mathematischen Argumentieren auch wahrnehmen. In den Vergleichsschulen nehmen die Schülerinnen und Schüler anfangs häufiger Gelegenheiten zum mathematischen Argumentieren wahr, die im Verlauf der Studie jedoch leicht abnehmen.

Ergebnisse

Tabelle 62. Deskriptive Statistiken zum Merkmal mathematisches Argumentieren, Schüler

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
mathematisches Argumentieren (Querschnitt)															
Projektgruppe	749	3.78 (1.03)				636	3.80 (1.06)				736	3.82 (1.10)			
Vergleichsgruppe	584	3.83 (1.05)		2.02	.00	457	3.92 (1.02)		3.66~	.00	564	3.95 (0.98)		4.70*	.00
gesamt	1333	3.79 (1.04)				1093	3.85 (1.04)				1300	3.87 (1.05)			
mathematisches Argumentieren (Längsschnitt)															
Projektgruppe	483	3.84 (0.99)				480	3.87 (1.07)				482	3.93 (1.06)			
Vergleichsgruppe	295	4.01 (0.91)		5.28*	.01	297	3.91 (0.99)		0.30	.00	298	3.95 (0.97)		0.06	.00
gesamt	778	3.91 (0.96)				777	3.88 (1.04)				780	3.94 (1.03)			
mathematisches Argumentieren (imputiert)															
Projektgruppe	718	3.74 (0.94)	.04			718	3.76 (1.01)	.04			718	3.83 (1.05)	.04		
Vergleichsgruppe	528	3.83 (0.88)	.04	2.80	.00	528	3.87 (0.98)	.04	3.60	.00	528	3.94 (0.93)	.04	3.59	.00
gesamt	1246	3.78 (0.92)	.03			1246	3.81 (1.00)	.03			1246	3.88 (1.00)	.03		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Im Merkmal Probleme lösen unterscheiden sich die Schülergruppen zu Beginn der Studie signifikant voneinander (Längsschnitt: $F(1, 776) = 8.13, p < .01$). Dabei haben die Schülerinnen und Schüler in den Vergleichsschulen signifikant häufiger die Wahrnehmung, Tätigkeiten des mathematischen Problemlösens angeboten zu bekommen als die Schülerschaft in den Projektschulen. Diese Unterschiede lösen sich zum zweiten Messzeitpunkt auf, indem in den Vergleichsschulen eine weniger häufigere Nachfrage berichtet wird, in den Projektschulen jedoch eine Steigerung stattgefunden hat. Die Projektschulen entwickeln sich zum dritten Messzeitpunkt auch noch weiter, die Vergleichsschulen erreichen wieder ihr Anfangsniveau. Eine zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor ergab einen sehr schwachen Haupteffekt für die Zeit ($F(2, 1500) = 2.49, p < .10$). Damit kann bestätigt werden, dass sich die Wahrnehmung aller Schülerinnen und Schüler in der Kompetenz Probleme lösen über die Zeit der Datenerhebung verändert hat. Der signifikante Mittelwertunterschied im Querschnitt zum dritten Messzeitpunkt dokumentiert einen Unterschied, in dem nach Wahrnehmung der Schülerinnen und Schüler in den Vergleichsschulen diese Tätigkeit dort häufiger eingefordert wird als in den Projektschulen ($F(1, 1298) = 13.90, p < .001$). Vergleicht man die Mittelwerte jedoch mit dem Datensatz des Längsschnitts, dann wird ersichtlich, dass zum dritten Messzeitpunkt viele Schülerinnen und Schüler an der Befragung teilnahmen, die den Gruppenmittelwert der Projektschulen nach unten verzerren. Diese unregelmäßige Teilnahme der Schülerinnen und Schüler könnte auf ein generelles Fehlzeitenproblem in der Schule hinweisen. Die unregelmäßige Teilnahme am Unterricht erschwert sodann die Reflexion über diesen.

Ergebnisse

Tabelle 63. Deskriptive Statistiken zum Merkmal Probleme mathematisch Lösen, Schüler

Merkmal	MZIP 1					MZIP 2					MZIP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
Probleme mathematisch Lösen (Querschnitt)															
Projektgruppe	748	3.62 (0.94)				636	3.67 (0.95)				736	3.66 (1.01)			
Vergleichsgruppe	583	3.74 (0.94)		5.67*	.00	457	3.75 (0.92)		2.09	.00	563	3.87 (0.89)		13.90***	.01
gesamt	1331	3.67 (0.94)				1093	3.71 (0.94)				1299	3.75 (0.97)			
Probleme mathematisch Lösen (Längsschnitt)															
Projektgruppe	482	3.69 (0.92)				480	3.73 (0.95)				482	3.76 (0.96)			
Vergleichsgruppe	295	3.87 (0.82)		8.13**	.01	297	3.78 (0.89)		0.55	.00	298	3.88 (0.86)		2.71	.00
gesamt	777	3.76 (0.89)				777	3.75 (0.93)				780	3.81 (0.92)			
Probleme mathematisch Lösen (imputiert)															
Projektgruppe	718	3.63 (0.86)	.03			718	3.64 (0.91)	.03			718	3.69 (0.95)	.04		
Vergleichsgruppe	528	3.75 (0.80)	.04	6.49	.01	528	3.72 (0.88)	.04	2.08	.00	528	3.86 (0.83)	.04	10.92	.01
gesamt	1246	3.68 (0.84)	.02			1246	3.67 (0.90)	.03			1246	3.76 (0.90)	.03		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Im Hinblick auf das Merkmal des mathematischen Modellierens wird auch hier beobachtet, dass die Schülerschaft der Vergleichsschulen eine häufigere Wahrnehmung aufweist, wenn es um mathematische Tätigkeiten geht, die diesem Merkmal zuzuordnen sind. Dieser Unterschied ist in der Betrachtung des Querschnitts noch statistisch signifikant, wird jedoch im Verlauf der Studie immer stärker abgeschwächt, indem die Projektschulen diese kompetenzorientierten Tätigkeiten häufiger einfordern. Dieser stetige Anstieg in den Modellierungsanforderungen kann auf den mit dem Projekt verbundenen Input für die Lehrkräfte zusammenhängen. Insgesamt zeigt sich auch, dass die Mittelwerte beider Gruppen am Ende jeweils über denen im Ausgangsniveau zu verorten sind. Für dieses Merkmal ist folglich insgesamt eine Entwicklung zu verzeichnen. Dieser Befund wird durch den Haupteffekt in einer Varianzanalyse mit Messwiederholung auf einem Faktor bestätigt ($F(2, 1502) = 4.97, p < .01$).

Ein Vergleich der grafischen Ergebnisse mit denen der Lehrkräfte zeigt jedoch die große Differenz in den Wahrnehmungen der beiden am Unterricht beteiligten Gruppen – Lehrkräfte und Schüler. Die Bemühungen der Lehrkräfte in den Projektschulen scheinen bei den Schülerinnen und Schülern nicht anzukommen, da sie sich ähnlich entwickeln wie die Gruppe der Vergleichsschulen. Ein Vergleich der Mittelwerte aus den Daten der Lehrkräfte und der Schülerinnen und Schüler in diesem Merkmal weist eine Differenz in den Skalenwerten von .89 (Projektschulen, Längsschnitt, 2. MZP) bzw. .87 (Projektschulen, Längsschnitt, 3. MZP) auf. Im Gegensatz dazu halbiert sich die Differenz nahezu in den Bewertungen der Vergleichsschulen⁶⁰. Gründe für diese unterschiedliche Wahrnehmung des Mathematikunterrichts könnten in einem veränderten Verständnis der Items auf Seiten der Lehrkräfte liegen. Ebenso wäre denkbar, dass die mathematischen Angebote der Lehrkräfte in den Projektschulen von den Schülerinnen und Schülern nicht wahrgenommen werden.

Gründe für die nahezu gleiche Entwicklung der Schülerstichproben könnten in der Verwendung des gleichen Arbeitsmaterials liegen. Alle Studienteilnehmer erhielten Anregungen durch die kompetenzorientierten Aufgaben des Leistungstests, für deren Lösung jeweils eine anzuwendende Kompetenz im Vordergrund steht. Dies bedeutet jedoch, dass die Lehrkräfte in den Projektschulen keine weiterführenden Anregungen aus der Projektarbeit ziehen konnten.

⁶⁰ Differenz Lehrkräfte (Längsschnitt, 2. MZP) und Schülerinnen und Schüler (Längsschnitt, 2. MZP): .47; Differenz im 3. MZP: .34 Skalenpunkte

Ergebnisse

Tabelle 64. Deskriptive Statistiken zum Merkmal mathematisch Modellieren, Schüler

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
mathematisches Modellieren (Querschnitt)															
Projektgruppe	747	3.28 (1.11)				635	3.29 (1.11)				736	3.40 (1.12)			
Vergleichsgruppe	583	3.44 (1.10)		7.03**	.01	457	3.43 (1.02)		4.00*	.00	565	3.51 (1.01)		3.29~	.00
gesamt	1330	3.35 (1.11)				1092	3.35 (1.08)				1301	3.45 (1.08)			
mathematisches Modellieren (Längsschnitt)															
Projektgruppe	482	3.34 (1.07)				479	3.33 (1.03)				482	3.48 (1.13)			
Vergleichsgruppe	295	3.43 (1.05)		1.32	.00	297	3.43 (1.03)		1.95	.00	298	3.52 (0.98)		0.29	.00
gesamt	777	3.38 (1.06)				776	3.37 (1.09)				780	3.49 (1.07)			
mathematisches Modellieren (imputiert)															
Projektgruppe	718	3.32 (1.02)	.04			718	3.30 (1.05)	.04			718	3.43 (1.08)	.04		
Vergleichsgruppe	528	3.44 (0.98)	.05	4.16	.00	528	3.43 (0.98)	.04	4.50	.00	528	3.51 (0.96)	.04	2.20	.00
gesamt	1246	3.37 (1.00)	.03			1246	3.35 (1.02)	.03			1246	3.46 (1.03)	.03		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Die durch die Lehrkräfte der Projektschulen im Verlauf häufiger initiierten Tätigkeiten zur Verwendung mathematischer Darstellungen werden auch von den Schülerinnen und Schülern wahrgenommen. Es zeigt sich zu Beginn der Studie, dass die Schülerinnen und Schüler in beiden Gruppen eine ähnliche Wahrnehmung hinsichtlich dieser mathematischen Kompetenz aufweisen. Eine signifikante Gruppenunterscheidung ist im ersten Messzeitpunkt nicht erkennbar. Zum zweiten Messzeitpunkt zeigt sich jedoch, dass diese Kompetenz in den Projektschulen sehr viel häufiger nachgefragt wird als in den Vergleichsschulen. Tendenziell geben beide Schülergruppen an, zum zweiten Messzeitpunkt häufiger mathematische Darstellungen zu verwenden. Ein Grund dafür könnte in einem inhaltlichen mathematischen Schwerpunkt der Jahrgangsstufe 9 und 10 liegen. Dennoch ergeben sich signifikante Unterschiede zu den einzelnen Messzeitpunkten. In den Projektschulen wird aus Sicht der Schülerinnen und Schüler diese Kompetenz signifikant häufiger nachgefragt als in den Vergleichsschulen. Diese Steigerung ist auch im Unterrichtsangebot der Lehrkräfte dokumentiert.

Auch zum dritten Messzeitpunkt unterscheiden sich die beiden Schülergruppen der Projekt- und Vergleichsschulen signifikant voneinander, wenngleich der Entwicklungsverlauf für beide Schülergruppen nahezu identisch erkennbar ist. Die Projektschüler nehmen eine Steigerung von 0.36 Skalenpunkte vom ersten zum zweiten Messzeitpunkt wahr (Längsschnitt), während die Schülerinnen und Schüler in den Vergleichsschulen im Mittel nur eine Steigerung von 0.20 Skalenpunkten dokumentierten. Beide halten dieses gesteigerte Niveau nahezu im dritten Messzeitpunkt. Mit Blick auf die Effektstärken kann jedoch nur von einem kleinen Effekt gesprochen werden ($\eta^2 = .02$; Längsschnitt, 2./3. MZP).

Ergebnisse

Tabelle 65. Deskriptive Statistiken zum Merkmal Darstellungen Verwenden, Schüler

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
Darstellungen Verwenden (Querschnitt)															
Projektgruppe	747	3.23 (0.92)				636	3.60 (0.96)				736	3.51 (0.99)			
Vergleichsgruppe	583	3.19 (0.94)		0.67	.00	457	3.39 (0.92)		13.46***	.01	565	3.33 (0.92)		11.41**	.01
gesamt	1330	3.21 (0.93)				1093	3.51 (0.95)				1301	3.43 (0.97)			
Darstellungen Verwenden (Längsschnitt)															
Projektgruppe	482	3.30 (0.90)				480	3.66 (0.95)				482	3.61 (0.98)			
Vergleichsgruppe	295	3.21 (0.92)		1.77	.00	297	3.41 (0.91)		13.09***	.02	298	3.33 (0.89)		16.55***	.02
Gesamt	777	3.27 (0.91)				777	3.56 (0.95)				780	3.50 (0.95)			
Darstellungen Verwenden (imputiert)															
Projektgruppe	718	3.28 (0.86)	.03			718	3.58 (0.93)	.04			718	3.54 (0.95)	.04		
Vergleichsgruppe	528	3.22 (0.85)	.04	1.14	.00	528	3.39 (0.89)	.04	14.08	.01	528	3.33 (0.88)	.04	15.75	.01
gesamt	1246	3.25 (0.86)	.03			1246	3.50 (0.92)	.03			1246	3.45 (0.93)	.03		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Bei der Betrachtung der mathematischen Kompetenz des technischen Arbeitens fällt auf, dass die Schülerinnen und Schüler der Vergleichsschulen zu Beginn der Studie angeben, signifikant häufiger aufgefordert zu werden, Tätigkeiten dieser Art durchzuführen. Im Vergleich mit den Angaben der Lehrkräfte fällt weiterhin auf, dass die Schülerinnen und Schüler diese Kompetenz auf sehr viel geringerem Niveau wahrnehmen. Mit Mittelwerten, die unter einem Skalenwert von vier liegen, haben sie viel weniger das Gefühl, technisch zu arbeiten bzw. Routinen einzuüben als es von den Lehrkräften initiiert wird. Im Verlauf der Studie haben beide Schülergruppen häufiger das Gefühl, zu kompetenzorientierten Tätigkeiten des technischen Arbeitens aufgefordert zu werden. Die Betrachtung im Längsschnitt zeigt, dass sich am Ende der Studie das Niveau in den Vergleichsschulen nicht verändert hat, die Projektschulen diese Kompetenz jedoch weiter ausbauten, bis auf das Niveau der Vergleichsschulen. Ein schwach signifikanter Unterschied zwischen den Gruppen ist am Ende lediglich im Querschnitt nachweisbar.

Mit Blick auf die Effektstärken unterscheiden sich die Gruppen jedoch nur gering. Eine zweifaktorielle Varianzanalyse mit Messwiederholung auf einem Faktor brachte keine signifikanten Ergebnisse hervor.

Ergebnisse

Tabelle 66. Deskriptive Statistiken zum Merkmal *technisch Arbeiten*, Schüler

Merkmal	MZP 1					MZP 2					MZP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
technisch Arbeiten (Querschnitt)															
Projektgruppe	749	3.82 (0.80)				636	3.94 (0.77)				735	3.96 (0.87)			
Vergleichsgruppe	583	3.91 (0.87)		4.00*	.00	457	4.10 (0.78)		10.68**	.01	565	4.05 (0.83)		3.95*	.00
gesamt	1332	3.86 (0.83)				1093	4.01 (0.78)				1300	4.00 (0.85)			
technisch Arbeiten (Längsschnitt)															
Projektgruppe	483	3.90 (0.78)				480	4.01 (0.75)				482	4.07 (0.81)			
Vergleichsgruppe	295	4.07 (0.74)		8.56**	.01	297	4.12 (0.78)		4.06*	.01	298	4.07 (0.82)		0.01	.00
gesamt	778	3.96 (0.77)				777	4.05 (0.77)				780	4.07 (0.81)			
technisch Arbeiten (imputiert)															
Projektgruppe	718	3.82 (0.74)	.03			718	3.92 (0.74)	.03			718	4.00 (0.78)	.03		
Vergleichsgruppe	528	3.93 (0.75)	.04	6.11	.00	528	4.04 (0.75)	.03	8.19	.01	528	4.04 (0.78)	.04	1.01	.00
gesamt	1246	3.87 (0.75)	.03			1246	3.97 (0.75)	.02			1246	4.02 (0.78)	.02		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke, Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Durchgängig signifikante Mittelwertunterschiede im Merkmal des mathematischen Kommunizierens ergeben sich nur im analysierten Datensatz des Querschnitts. Die Vergleichsschulen berichten hier durchgängig einen höheren Skalenwert und damit eine häufigere Nachfrage nach Tätigkeiten des mathematischen Kommunizierens. Die Mittelwerte in den Projektschulen verändern sich nicht, während in den Vergleichsschulen zumindest kleine Veränderungen berichtet werden. Die Schülerinnen und Schüler, die im Datensatz des Querschnitts für geringere Mittelwerte in den Projektschulen sorgten, haben jedoch nicht durchgängig am Projekt teilgenommen. Ein Vergleich mit dem Längsschnitt zeigt, dass hier die Schülerinnen und Schüler von Anfang an höhere Werte berichten, die jedoch noch minimal abfallen. Eine von den Schülerinnen und Schülern häufiger wahrgenommene Unterrichtsbeteiligung in Form des mathematischen Kommunizierens konnte in den Projektschulen nicht erreicht werden.

Mit Blick auf die signifikanten Gruppenunterschiede zum Beginn der Studie im Merkmal mathematisch Kommunizieren, kann ebenso wie im mathematischen Argumentieren und Probleme lösen sowie beim technischen Arbeiten von einer Negativauswahl der Schülerinnen und Schüler in den Projektschulen ausgegangen werden.

Ergebnisse

Tabelle 67. Deskriptive Statistiken zum Merkmal mathematisch Kommunizieren, Schüler

Merkmal	MZIP 1					MZIP 2					MZIP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
mathematisch Kommunizieren (Querschnitt)															
Projektgruppe	749	3.78 (1.00)				636	3.78 (1.04)				735	3.76 (1.05)			
Vergleichsgruppe	583	3.92 (1.03)		6.30*	.01	457	4.00 (0.95)		13.27***	.01	565	3.97 (0.99)		14.04***	.01
gesamt	1332	3.84 (1.02)				1093	3.87 (1.01)				1300	3.85 (1.03)			
mathematisch Kommunizieren (Längsschnitt)															
Projektgruppe	483	3.86 (0.98)				480	3.82 (1.03)				482	3.83 (1.02)			
Vergleichsgruppe	295	4.05 (0.92)		6.91**	.01	297	3.97 (0.94)		3.84~	.01	298	3.95 (0.98)		2.32	.00
gesamt	778	3.93 (0.96)				777	3.88 (1.00)				780	3.88 (1.00)			
mathematisch Kommunizieren (imputiert)															
Projektgruppe	718	3.78 (0.93)	0.035			718	3.75 (0.99)	0.037			718	3.78 (0.99)	0.037		
Vergleichsgruppe	528	3.92 (0.89)	0.04	7.28	.01	528	3.93 (0.91)	0.041	10.96	.01	528	3.96 (0.94)	0.042	11.09	.01
gesamt	1246	3.84 (0.92)	0.026			1246	3.83 (0.96)	0.028			1246	3.85 (0.97)	0.028		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (MW), Standardabweichung (SD), Standardfehler des Mittelwertes (SE), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (N); Angaben aus dem imputierten Datensatz entsprechen dem gepoolten Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt, naives Pooling)

Tabelle 68 bestätigt durch die finalen Wachstumsmodelle unter Berücksichtigung der unterschiedlichen Zeitpunkte der Datenerhebungen noch einmal übersichtlich die bereits beschriebenen Ergebnisse: Der Unterricht in den Projektschulen unterscheidet sich aus Sicht der Schülerinnen und Schüler signifikant vom Unterricht in den Vergleichsschulen in den Merkmalen des mathematischen Argumentierens, Probleme lösens, technisch arbeiten und Kommunizieren. Eine erkennbare und statistisch nachweisbare Veränderung des Unterrichts wird von den Schülerinnen und Schülern der Projektschulen jedoch nur im Merkmal der Verwendung von Darstellungen dokumentiert. Für alle anderen Unterrichtsmerkmale ist die Entwicklung zwar deskriptiv sichtbar, statistisch aber nicht haltbar. Ein Blick auf die zufälligen Effekte zeigt, dass für den durchschnittlichen Schüler keine erklärbare Varianz in seiner individuellen Wachstumskurve verbleibt. Allerdings weisen die signifikanten Varianzkomponenten zwischen den Schülern darauf hin, dass es noch weitere Faktoren neben der durchgeführten Intervention geben muss, die sowohl den Aufgangswert als auch die Wachstumsrate erklären können.

Tabelle 68. Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 1, Schülerinnen und Schüler

		Parameter	Argumentieren	Probleme Lösen	Modellieren	Darstellungen	technisch Arbeiten	Kommunizieren
<i>Feste Effekte</i>								
Ausgangswert	Konstante	γ_{00}	3.816*** (.046)	3.721*** (.038)	3.417*** (.045)	3.241*** (.041)	3.930*** (.036)	3.915*** (.041)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	-.088 (.057)	-.095~ (.050)	-.120* (.058)	.060 (.050)	-.121** (.044)	-.142** (.054)
Wachstumswert	Konstante	γ_{10}	.008~ (.004)	.006~ (.004)	.005 (.004)	.010* (.004)	.010** (.004)	.003 (.004)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	-.001 (.005)	-.003 (.005)	.002 (.005)	.012* (.005)	.004 (.004)	-.003 (.005)
<i>Zufallseffekte</i>								
	innerhalb Personen	σ^2_{ε}	.483	.437	.561	.440	.305	.480
Level 1	Ausgangswert	σ^2_0	.367***	.273***	.450***	.298***	.250***	.370***
Level 2	Wachstumswert	σ^2_1	.001***	.001***	.002***	.001***	.001***	.001***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

6. 3.1.5 Entwicklung der Mathematikleistung

Die Leistungen der Schülerinnen und Schüler in der neunten Jahrgangsstufe und damit am Beginn der Studie unterscheiden sich signifikant voneinander. Die Schülerinnen und Schüler in den Projektschulen weisen mit einem Mittelwert von $MW=502.16$ signifikant höhere Leistungswerte auf als die Schülerinnen und Schüler in den Vergleichsschulen ($MW=468.29$). Die Vergleichsgruppe stellt damit eine Negativauswahl an Lernerinnen und Lernern dar ($F(1, 1348) = 43.9, p < .001$). Zur Einschätzung der vorliegenden Leistungen wurde ein Vergleich mit den globalen Kompetenzwerten der betreffenden Bundesländer Berlin und Brandenburg aus dem IQB-Ländervergleich von 2012 vorgenommen (Roppelt et al., 2013). In dieser Studie ergaben sich repräsentative, globale Mittelwerte für die einzelnen Bundesländer. Mit einem mittleren Kompetenzwert von $MW(BB) = 518$ Punkten zeigen die Brandenburger Schülerinnen und Schüler im Ländervergleich bessere Mathematikleistungen als die Berliner Lerner mit einem Wert von $MW(BE) = 479$. Tabelle 69 zeigt die Abweichungen der Mittelwerte für die vorliegende Studie von diesem repräsentativen Globalwert im Fach Mathematik im jeweiligen Bundesland.

Tabelle 69. Gegenüberstellung der Mittelwerte aus den teilnehmenden Schulen mit den Mittelwerten der Bundesländer aus dem IQB-Ländervergleich 2012

Gruppe	Berlin (BE)		Brandenburg (BB)	
	Mittelwert Studie 1	Globalwert IQB- LV	Mittelwert Studie 1	Globalwert IQB- LV
Projektschulen	515.66	479.00	481.24	518.00
Vergleichsschulen	419.46	479.00	514.27	518.00

Anmerkung. Globalwert IQB-LV (Ländervergleich) als Mittel der Kompetenzstände der Schülerinnen und Schüler in der 9. Jahrgangsstufe

Mit einem angenommenen Lernzuwachs von 20-30 Kompetenzpunkten pro Schuljahr zeigt sich, dass die Berliner Projektschulen die Brandenburger Projektschulen um mehr als ein Lernjahr übersteigen. Weiterhin weichen die Berliner Schüler der Projektschulen erheblich vom Landesmittelwert ab, während in Brandenburg eine Negativauswahl an Schülerinnen und Schülern für die Projektschulen vorliegt. Deren Werte liegen ca. ein Lernjahr hinter dem brandenburgischen Mittelwert.

Bei den Vergleichsschulen zeigt sich, dass die Leistungen der teilnehmenden Schulen im Land Brandenburg dem Landesmittelwert entspricht. Im Bundesland Berlin ist jedoch die große negative Abweichung der Mathematikleistungen vom Landesmittelwert erkennbar. Die teilnehmenden Berliner Schülerinnen und Schüler in den Vergleichsschulen stellen

eine erhebliche Negativauswahl dar. Ihre Leistungen liegen um nahezu zwei Lernjahre hinter dem Landesdurchschnitt und um ca. 3-4 Lernjahre hinter den Brandenburgern. Ursache für die geringen mittleren Leistungswerte der teilnehmenden Schülerinnen und Schüler sind somit die Lerner in den Vergleichsschulen im Bundesland Berlin.

Mit Blick auf Tabelle 70 ist jedoch vom ersten zum zweiten ein sehr großer Kompetenzzuwachs zu verzeichnen, der jedoch nicht nachhaltig ausgebaut werden kann, die Kompetenzwerte im dritten Messzeitpunkt fallen gegenüber dem zweiten Messzeitpunkt wieder leicht ab. Der signifikante Unterschied zwischen den beiden Gruppen bleibt über die gesamte Studiendauer erhalten. In einer zweifaktoriellen Varianzanalyse mit Messwiederholung auf einem Faktor wird die Veränderung der Mathematikleistung durch einen Haupteffekt bestätigt ($F(2,1520)=236.73, p<.001$). Ein weiterer Haupteffekt ist in der Veränderung der beiden Gruppen über die Zeit zu beobachten ($F(1,760)=21.00, p<.001$). Eine unterschiedliche Entwicklung der Mathematikleistung der beiden Gruppen in Form eines Interaktionseffektes kann jedoch nicht nachgewiesen werden (Abbildung 46).

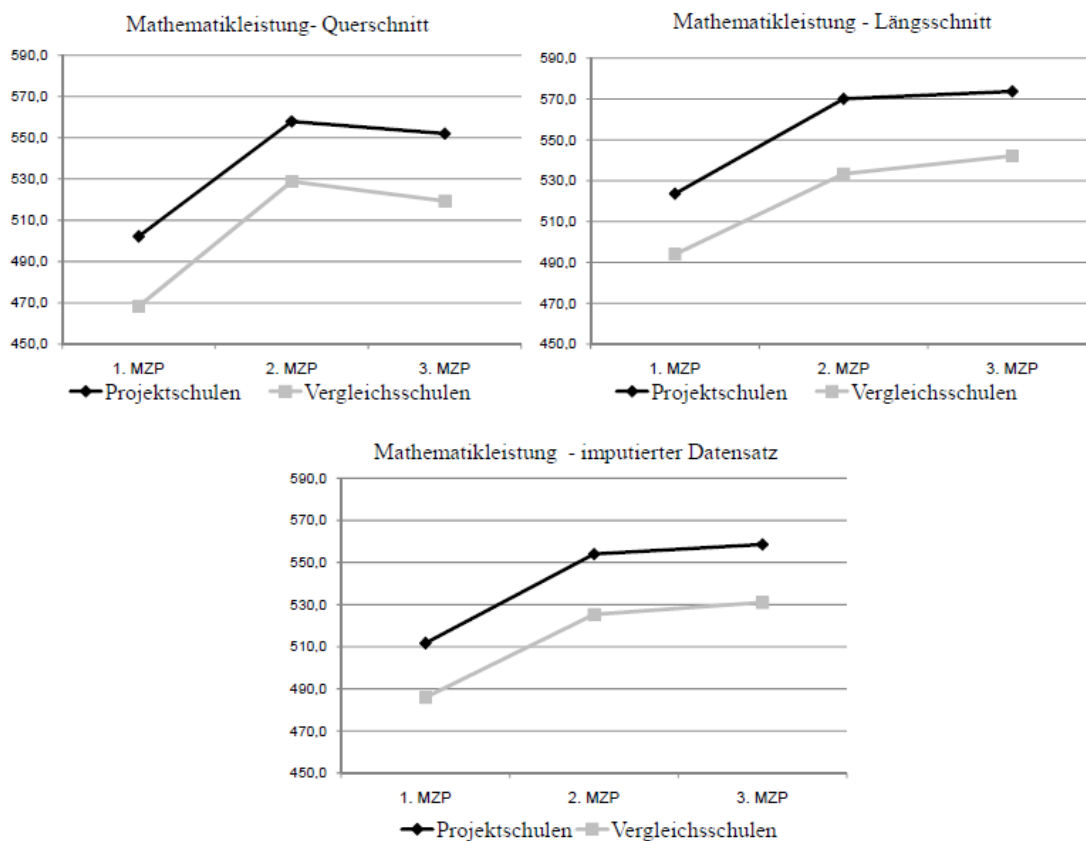


Abbildung 46. Mathematikleistung im Verlauf - Gegenüberstellung aller Teilnehmer, Längsschnitt und imputierter Datensatz.

Ergebnisse

Tabelle 70. Deskriptive Statistiken zum Merkmal der Mathematikleistung, Schüler

Merkmal	MZIP 1					MZIP 2					MZIP 3				
	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SD)</i>	<i>SE</i>	<i>F</i>	η^2
Mathematikleistung (Querschnitt)															
Projektgruppe	747	502.16 (95.02)				643	557.91 (105.30)				740	552.04 (109.07)			
Vergleichsgruppe	602	468.29 (91.16)		43.92***	.03	462	528.80 (99.17)		21.56***	.02	581	519.33 (110.70)		28.90***	.02
gesamt	1349	487.04 (94.80)				1105	545.74 (103.73)				1299	537.65 (110.95)			
Mathematikleistung (Längsschnitt)															
Projektgruppe	476	523.61 (86.61)				482	570.13 (102.86)				486	573.76 (103.96)			
Vergleichsgruppe	303	493.98 (92.05)		20.63***	.03	298	533.40 (100.84)		23.82***	.03	299	542.18 (106.80)		16.72***	.02
gesamt	779	512.09 (89.88)				779	556.10 (103.58)				785	561.73 (106.10)			
Mathematikleistung (imputiert)															
Projektgruppe	719	511.65 (90.45)	3.47			719	554.00 (107.01)	4.00			719	558.59 (106.40)	4.07		
Vergleichsgruppe	529	485.92 (89.40)	4.23	24.95	.02	529	525.32 (99.32)	4.45	23.28	.02	529	530.95 (104.37)	4.60	20.91	.02
gesamt	1248	500.74 (90.86)	2.77			1248	541.84 (104.75)	3.00			1248	546.88 (106.39)	3.11		

Anmerkungen. ~ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$, Mittelwert (*MW*), Standardabweichung (*SD*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*); Angaben aus dem imputierten Datensatz entsprechen dem Mittel aus fünf Imputationen (Signifikanzen nicht angezeigt)

Die Wachstumsmodelle bestätigen die bisherigen Ergebnisse. Stellt man die unterschiedlichen Messzeitpunkte in Rechnung, dann weist ein durchschnittlicher Schüler in den Vergleichsschulen einen mathematischen Kompetenzwert von 486.67 aus. Ein Schüler, der in den Projektschulen unterrichtet wurde, weist einen durchschnittlichen Kompetenzwert von 514.08 aus. Der monatliche Kompetenzzuwachs eines Schülers in den Vergleichsschulen liegt bei 3.7 Kompetenzpunkten. Das monatliche Kompetenzwachstum der Schülerinnen und Schüler in den Projektschulen weist einen Wert von 3.081 auf, der jedoch als statistisch unbedeutend zu bewerten ist und somit im Bereich des Zufälligen liegt. In ihrem Lernzuwachs unterscheiden sich die Schülerinnen und Schüler in den Projektschulen damit nicht von den Schülerinnen und Schülern in den Vergleichsschulen. Die von den Lehrkräften in den Projektschulen berichtete kompetenzorientierte Unterrichtsgestaltung wird von den Schülerinnen und Schülern nur durch die vermehrte Verwendung von mathematischen Darstellungen wahrgenommen (Abschnitt 6.3.1.4) und schlägt sich noch nicht in den Leistungen bei der Bearbeitung kompetenzorientierter Aufgaben nieder.

Tabelle 71. Wachstumseffekte in der Mathematikleistung der Schülerinnen und Schüler, Studie 1

		Parameter	Leistung
<i>Feste Effekte</i>			
Ausgangswert	Konstante	γ_{00}	486.674*** (4.162)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	27.410*** (5.342)
Wachstumswert	Konstante	γ_{10}	3.731*** (.256)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	-.065 (.306)
<i>Zufallseffekte</i>			
	innerhalb Personen	σ^2_{ϵ}	2062.487
Level 1	Ausgangswert	σ^2_0	6171.876***
Level 2	Wachstumswert	σ^2_1	2.240

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

Die Zufallseffekte weisen darauf hin, dass die Varianz zwischen den Messzeitpunkten bzw. innerhalb der Personen durch das Wachstumsmodell vollständig erklärt wurden. Ebenfalls verbleibt keine erklärbare Varianz auf dem Wachstumswert. Einzig die Varianz

im Ausgangswert (durchschnittliche Mathematikleistung zum Messzeitpunkt Null) zwischen den Schülerinnen und Schülern (Ebene 2) weist darauf hin, dass andere Faktoren zur Erklärung herangezogen werden sollten. Welche Merkmale die Mathematikleistung beeinflussen, wurde in verschiedenen Studien bereits intensiv analysiert⁶¹. Interessant erscheint jedoch der Vergleich der Wachstumsmodelle zur Wahrnehmung des Unterrichts (Tabelle 68) und zur Mathematikleistung (Tabelle 71). Für die Ausgangswerte kann formuliert werden, dass ein geringerer Grad an Kompetenzorientierung in den Projektschulen dennoch höhere Leistungen bei der Lösung kompetenzorientierter Aufgaben hervorbrachte. Die Wachstumswerte zeigen nur im Bereich des technischen Arbeitens einen signifikanten Zuwachs bei den Projektschulen auf. In den anderen mathematischen Kompetenzen sind es eher die Vergleichsschulen, die ein schwach signifikantes Wachstum zeigen. Das zeigt sich auch im signifikanten Leistungszuwachs der Vergleichsschulen. Ein gemeinsames Modell von Unterrichts- und Leistungsvariablen wäre hier wünschenswert, würde jedoch von einem Zweiebenenmodell schnell zu einem Drei- bzw. Vierebenenmodell führen.

6.3.2 Studie 2 „Implementation der Bildungsstandards durch unterstützte Unterrichtsentwicklung“ – Ergebnisse aus Schülersicht

Studie 2 ermöglicht es, die Unterrichtsentwicklungskonzepte der Bundesländer Berlin/Brandenburg und Hessen miteinander zu vergleichen und damit das vom IQB entwickelte Konzept zu evaluieren. In der Befragung der Lehrkräfte (Abschnitt 6.2.2) zeigte sich bereits, dass sich das Unterrichtsangebot der Lehrkräfte in Hessen nur wenig veränderte. Die Lehrkräfte in den Bundesländern Berlin/Brandenburg berichteten jedoch für nahezu alle kompetenzorientierten mathematischen Tätigkeiten eine häufigere Initiierung in ihrem Unterricht. Eine Analyse der Schülerwahrnehmungen aus den einzelnen Bundesländern führte zu keinen signifikanten Ergebnissen, die sich auf die veränderte Unterrichtsgestaltung zurückführen lassen. Vielmehr kann insbesondere durch die grafische Darstellung gezeigt werden, dass die Schülerinnen und Schüler im Bundesland Hessen eine gemeinsame Entwicklung des Unterrichts wahrnehmen. Die anfänglichen Unterschiede zwischen den Projekt- und Vergleichsschulen bleiben mehrheitlich bestehen. Die Veränderung zwischen den beiden Messzeitpunkten zeigt nahezu gleiche Wachstumsko-

⁶¹ Allgemeine Einflussfaktoren auf die Schülerleistung in Hattie (2009); speziell zur Mathematikleistung vgl. Abschnitt 2.3.2

effizienten auf, im Gegensatz zur Schülerwahrnehmung in den Bundesländern Berlin/Brandenburg. Insgesamt weisen die Schülerinnen und Schüler aus Berlin/Brandenburg ein in etwa gleich hoch ausgeprägtes mathematisches Selbstkonzept und einen gleich hohen Grad an mathematischem Interesse auf wie die Mitschüler in Hessen. Die Schülergruppen unterscheiden sich in den einzelnen Messzeitpunkten nicht voneinander.

Tabelle 72. Deskriptive Statistiken zum Merkmal des mathematischen Selbstkonzeptes, Schüler, Studie 2

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>math. Selbstkonzept (global)</i>								
Projektgruppe	1.082	2.48 (.023)	.43	.000	1.079	2.50 (.024)	.07	.00
Vergleichsgruppe	854	2.50 (.027)			851	2.51 (.028)		
<i>math. Selbstkonzept (BB/BB)</i>								
Projektgruppe	382	2.42 (.037)			372	2.41 (.038)		
Vergleichsgruppe	380	2.48 (.039)			371	2.49 (.041)		
<i>math. Selbstkonzept (Hessen)</i>								
Projektgruppe	700	2.51 (.030)	1.15	.00	707	2.55 (.031)	2.67*	.00
Vergleichsgruppe	474	2.51 (.038)			480	2.53 (.038)		
gesamt	1.936	2.49 (.018)			1.930	2.51 (.018)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Während die Gruppenmittelwerte im Merkmal des mathematischen Selbstkonzeptes um den theoretischen Mittelwert schwanken, liegen diese beim Merkmal des mathematischen Interesses noch darunter. Die Schülerinnen und Schüler konnten im Verlauf ihrer schulischen Laufbahn kein besonderes Interesse an dem Fach entwickeln. Tabelle 73 weist Mittelwerte von 2.04 bis maximal 2.16 Skalenpunkten aus.

Ergebnisse

Tabelle 73. Deskriptive Statistiken zum Merkmal des mathematischen Interesses, Schüler, Studie 2

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>math. Interesse (global)</i>								
Projektgruppe	1.081	2.08 (.021)	2.44	.00	1.076	2.06 (.022)	2.19	.00
Vergleichsgruppe	853	2.13 (.025)			851	2.11 (.025)		
<i>math. Interesse (BB/BB)</i>								
Projektgruppe	382	2.06 (.035)			370	2.04 (.036)		
Vergleichsgruppe	380	2.16 (.039)			371	2.15 (.039)		
<i>math. Interesse (Hessen)</i>								
Projektgruppe	699	2.09 (.027)	1.31	.00	706	2.08 (.027)	1.67	.00
Vergleichsgruppe	473	2.11 (.033)			480	2.08 (.032)		
gesamt	1.934	2.10 (.016)			1.927	2.08 (.016)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Einzig die Lehrerbewertung wird in Hessen positiver bewertet als in Berlin/Brandenburg. Während in den Hessischen Projektschulen eine Steigerung von .31 Skalenpunkten zu beobachten ist, stagnieren die Schülereinschätzungen in Berlin/Brandenburg (Tabelle 74). Eine durch die Schülerinnen und Schüler wahrnehmbare Änderung im Unterrichtsverhalten der Lehrkräfte in BB/BE kann in den Projektschulen demnach nicht bestätigt werden.

Tabelle 74. Deskriptive Statistiken zum Merkmal der Lehrerbewertung, Schüler, Studie 2

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>Lehrerbewertung (global)</i>								
Projektgruppe	1.075	2.90 (.025)	.04	.00	1.072	2.93 (.024)	1.24	.00
Vergleichsgruppe	850	2.89 (.027)			849	2.97 (.027)		
<i>Lehrerbewertung (BB/BB)</i>								
Projektgruppe	377	2.70 (.042)	19.01***	.03	367	2.70 (.042)	24.74***	.04
Vergleichsgruppe	379	2.75 (.040)			371	2.84 (.041)		
<i>Lehrerbewertung (Hessen)</i>								
Projektgruppe	698	2.75 (.040)	19.01***	.03	705	3.06 (.027)	24.74***	.04
Vergleichsgruppe	471	3.00 (.037)			478	3.08 (.036)		
gesamt	1.925	2.89 (.018)			1.921	2.95 (.018)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Stellt man die unterschiedlichen Messzeitpunkte in Rechnung, dann steigert ein durchschnittlicher Schüler aus dem Bundesland Berlin/Brandenburg in den Vergleichsschulen

seine Lehrerbewertung um monatlich .01 Skalenpunkte. Dieses schwache Wachstum kann durch die Itemformulierungen erklärt werden. Die aus sechs Items bestehende Skala der Lehrerbewertung zielt mit drei Items auf eine Einschätzung der Sympathie zwischen Lehrkraft und Schüler und mit den verbleibenden drei Items auf die Ausstrahlung der Lehrkraft bzw. die Begeisterung der Lehrkraft für Mathematik ab. Für Schülerinnen und Schüler spielt bei der Bewertung der zwischenmenschlichen Beziehung andere Faktoren eine Rolle als die Unterrichtsgestaltung der Lehrkraft. Oftmals werden in Items wie „Ich bin mit meinem Mathematiklehrer/meiner Mathematiklehrerin sehr zufrieden.“ die Gerechtigkeit bei Notengebung, die dem Schüler entgegengebrachte Aufmerksamkeit der Lehrkraft, die Fähigkeit zum Erklären schwieriger Sachverhalte und das eigene Erfolgserleben im Unterricht mitbewertet. Vor diesem Hintergrund ist die Bewertung der Lehrkraft in den Hessischen Projektschulen zum zweiten Messzeitpunkt gegenüber dem ersten Messzeitpunkt erstaunlich hoch. Die Lehrkraft wurde im zweiten Messzeitpunkt um .31 Skalenpunkte höher bewertet. In den anderen Gruppen erreichte die Lehrerbewertung eine maximale Steigerung von .09 Skalenpunkten. Gründe für die Schülerbewertung in den Hessischen Projektschulen sind nur schwer auszumachen und höchst spekulativ.

Tabelle 75. Wachstumseffekte zum mathematischen Selbstkonzept, Interesse an Mathematik und der Lehrerbewertung der Schülerinnen und Schüler, Studie 2

		Parameter	math. Selbstkonzept	Interesse	Lehrerbewertung
<i>Feste Effekte</i>					
Ausgangswert	Konstante	γ_{00}	2.479*** (.040)	2.138*** (.043)	2.687*** (.053)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	-.037 (.040)	-.057 (.039)	-.007 (.047)
	Bundesland (0=BE/BB, 1=HE)	γ_{02}	.037 (.042)	-.009 (.043)	.290*** (.056)
Wachstums- wert	Konstante	γ_{10}	-.003 (.004)	.002 (.005)	.011~ (.006)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	.002 (.003)	.001 (.004)	-.006 (.005)
	Bundesland (0=BE/BB, 1=HE)	γ_{12}	.006 (.004)	.000 (.005)	.001 (.007)
<i>Zufallseffekte</i>					
	innerhalb Personen	σ^2_{ε}	.138	.186	.353
Level 1	Ausgangswert	σ^2_0	.488***	.326***	.250***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, BE/BB = Berlin/Brandenburg, HE = Hessen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparameters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

Die Schülerbewertung zur kompetenzorientierten Unterrichtsgestaltung durch die Lehrkräfte wird grundsätzlich von den Hessischen Schülerinnen und Schülern positiver bewertet. Für die Kompetenz des mathematischen Argumentierens zeichnet sich im Hessischen Unterricht eine Häufigkeit ab, die deutlich über der des Unterrichts in Berlin/Brandenburg liegt. Bereits in der Ausgangserhebung unterscheiden sich sowohl die Bundesländer als auch die Gruppen (Projekt- und Vergleichsschulen) voneinander. Dabei wird in den Projektschulen am Beginn des Schuljahres häufiger mathematisch argumentiert als in den Vergleichsschulen. Eine signifikante Änderung bis zum Schuljahresende ist jedoch nicht zu beobachten. Bei globaler Betrachtung der Projekt- und Vergleichsschulen – unabhängig vom Bundesland – nähern sich die Gruppen an (Anhang E). Das deutet darauf hin, dass die Lehrkräfte in den Projektschulen die Unterstützung und Möglichkeiten des Implementationsprojektes nicht im Unterricht umsetzen konnten. Für die Hessischen Lehrkräfte in den Projektschulen decken sich diese Ergebnisse nicht mit ihrer Selbstauskunft. In Abschnitt 6.2.2 geben die Lehrkräfte an, diese Kompetenz nicht gefördert zu haben. Allerdings nehmen die Hessischen Schülerinnen und Schüler eine andere Entwicklung wahr. Ihrer Auskunft nach wurde im Unterricht häufiger von ihnen verlangt, mathematisch zu argumentieren. Eine ähnliche Divergenz zeigt sich für die Lehrkräfte in den Projektschulen in Berlin/Brandenburg. Sie geben an, diese Kompetenz häufiger eingefordert zu haben (Studie 1, Abbildung 40), die Schülerinnen und Schüler nehmen jedoch eine gegenteilige Entwicklung wahr. In den Vergleichsschulen verlaufen die Unterrichtswahrnehmungen jedoch richtungsgleich von Lehrkräften und Schülern.

Eine Varianzanalyse mit Messwiederholung bestätigt statistisch die Unterschiedlichkeit der Gruppen. Ein anschließender Gruppenvergleich zeigt, dass sich die Projekt- und Vergleichsschulen in Berlin/Brandenburg nicht signifikant unterschiedlich entwickeln. Mit der Referenzgruppe der Vergleichsschulen aus Berlin/Brandenburg zeigen sich jedoch deutliche Signifikanzen in der Entwicklung der hessischen Projekt- und Vergleichsschulen.

Im Bereich des Probleme mathematisch lösen geben die Schülerinnen und Schüler der Projektschulen zunächst an, gegenüber der jeweiligen Vergleichsgruppe im Bundesland häufiger zur Anwendung dieser Kompetenz aufgefordert zu werden. Im Verlauf eines Schuljahres sind es jedoch die Schülerinnen und Schüler in den Vergleichsschulen, die

eine leichte Steigerung wahrnahmen. In den Projektschulen können die Schülerinnen und Schüler eine häufigere Initiierung von Tätigkeiten zum Probleme lösen nur in Hessen bestätigen. In BB/BE liegt erneut eine große Divergenz in der Schüler- und Lehrerwahrnehmung in den Projektschulen vor (Abbildung 41). Das Unterrichtsangebot im Bereich des Probleme lösens wird von den Schülerinnen und Schülern in den Projektschulen in BE/BB gegenteilig bewertet. Wenngleich also für das IQB-Implementationskonzept aus Sicht der Lehrkräfte in den Projektschulen ein positives Ergebnis erzielt wird, relativieren die Schülerinnen und Schüler als ebenfalls Beteiligte am Unterricht dieses Ergebnis.

Unter Einbezug des evaluativen Aspektes der beiden Implementationskonzepte zeigt sich, dass die Schülerinnen und Schüler in Berlin/Brandenburg eine Veränderung des Unterrichts wahrnehmen. In den Vergleichsschulen erhöht sich die Häufigkeit dieser mathematischen Kompetenz in nahezu gleichem Maße in Hessen und BB/BE. Auch die Projektschulen in Hessen weisen aus Sicht der Schülerinnen und Schüler eine leichte Steigerung auf. Die Unterrichtswahrnehmung der Lehrkräfte und Schülerinnen und Schüler deckt sich.

Für den bei deutschen Schülerinnen und Schülern unterentwickelten Bereich des Modellierens zeigt sich zunächst Einigkeit in der Frage, wie häufig Tätigkeiten dieser Kompetenz im Unterricht initiiert werden. Die Gruppen und Bundesländer unterscheiden sich nicht voneinander. Im Verlauf des Schuljahres sind es lediglich die Hessischen Schülerinnen und Schüler, die eine Entwicklung wahrnehmen. Die Schülerinnen und Schüler in den Projektschulen in Berlin/Brandenburg nehmen sogar eine Abnahme der Tätigkeiten wahr, die mathematisches Modellieren verlangen. Hier unterscheiden sich Schüler- und Lehrersicht erneut ganz entscheidend voneinander. Die Lehrkräfte gaben an, gerade diese mathematische Kompetenz besonders zu fördern.

Aus Schülersicht nähern sich die Wahrnehmungen in der Kompetenz mathematische Darstellungen verwenden im Laufe eines Schuljahres an. Am Ende zeigt sich ein einheitliches Bild bzw. gleiche Wahrnehmungen in allen Gruppen und beiden Regionen im Hinblick auf diese Kompetenz. Statistisch steigern sich die Projektschulen signifikant in ihrer Entwicklung und schließen damit an die Wahrnehmungen der Schüler aus den Vergleichsschulen auf.

Das Abarbeiten von mathematischen Routinen wird im Verlauf eines Schuljahres in beiden Gruppen weiter gefördert. Es ist ein signifikantes Wachstum zu beobachten. In Hessen nehmen die Schülerinnen und Schüler das routinierte Abarbeiten von mathematischen

Algorithmen signifikant häufiger wahr als die Schüler in den Bundesländern Berlin/Brandenburg. Im differenzierten Blick wird deutlich, dass hessische Schülerinnen und Schüler häufiger diese Kompetenz anwenden als die Schüler in Berlin/Brandenburg. In den Projektschulen wird in Berlin/Brandenburg keine weitere Ausdehnung dieser Kompetenz wahrgenommen, während in allen anderen Gruppen die Häufigkeit zur Anwendung technischer Rechenfertigkeiten steigt. Die Mehrebenenmodelle weisen eine statistische Signifikanz für die Entwicklung in den hessischen Schulen dahingehend auf, dass das Wachstum weniger stark ausgeprägt und damit im *Sinne* eines gewünschten Mathematikunterrichts ausfällt als in den restlichen Schulen (Tabelle 76).

Im Gesamtblick auf den Unterricht bleibt festzuhalten, dass es für die Ausprägung der Kompetenzorientierung einen Unterschied macht, in welcher Schule die Schülerinnen und Schüler unterrichtet werden und auch, in welchem Bundesland die Schülerinnen und Schüler wohnhaft sind. Ein unterstützendes Unterrichtsentwicklungskonzept, dass das Unterrichtsangebot durch die Lehrkräfte verändert, wird jedoch nur punktuell wahrgenommen. Diese Wahrnehmung der Schülerinnen und Schüler unterscheidet sich zusätzlich stark von der Intention der Lehrkräfte.

Ergebnisse

Tabelle 76. Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 2, Schülerinnen und Schüler

		Parameter	Argumentieren	Probleme Lösen	Modellieren	Darstellungen	technisch Arbeiten	Kommunizieren
<i>Feste Effekte</i>								
Ausgangswert	Konstante	γ_{00}	3.506*** (.065)	3.478*** (.064)	3.372*** (.069)	3.465*** (.063)	3.604*** (.058)	3.702*** (.065)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	.128* (.059)	.119* (.055)	.030 (.065)	-.197** (.056)	.037 (.049)	.031 (.059)
	Bundesland (0=BE/BB, 1=HE)	γ_{02}	.403*** (.067)	.316*** (.065)	.004 (.072)	.050 (.063)	.320*** (.059)	.212 (.067)
Wachstumswert	Konstante	γ_{10}	.012 (.008)	.014~ (.008)	-.004 (.008)	-.007 (.008)	.033*** (.007)	.010 (.008)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	-.007 (.007)	-.014* (.006)	.002 (.007)	.017* (.007)	-.010 (.006)	-.010 (.007)
	Bundesland (0=BE/BB, 1=HE)	γ_{12}	.012 (.008)	.006 (.008)	.013 (.009)	.001 (.008)	-.013~ (.007)	.011 (.008)
<i>Zufallseffekte</i>								
	innerhalb Personen	σ^2_{ε}	.546	.521	.666	.588	.447	.557
Level 1	Ausgangswert	σ^2_0	.369***	.322***	.467***	.240***	.196***	.376***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, BE/BB = Berlin/Brandenburg, HE = Hessen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparameters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

Hinsichtlich der Mathematikleistung lässt sich durch Mittelwertvergleiche nachweisen, dass sich die Projekt- und Vergleichsschulen in ihren Ausgangswerten signifikant voneinander unterscheiden, diese Bedeutsamkeit jedoch auf die unterschiedlichen Bundesländer zurückzuführen ist (Tabelle 78). Ein differenzierter Blick zeigt auf, dass die Schülerinnen und Schüler aus Berlin/Brandenburg in ihren Mathematikleistungen stark hinter den hessischen Lernern zurückliegen. In dieser deskriptiven Ansicht der Daten bleibt jedoch der unterschiedliche Testzeitraum unberücksichtigt. Eine Vielzahl von hessischen Schülerinnen und Schülern befand sich zum ersten Messzeitpunkt bereits in der zehnten Jahrgangsstufe, wodurch die enormen Leistungsunterschiede erklärt werden können.

Tabelle 77. Deskriptive Statistiken zur Mathematikleistung, Schüler

Merkmal	MZP 1				MZP 2			
	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2	<i>N</i>	<i>MW (SE)</i>	<i>F</i>	η^2
<i>Mathematikleistung (global)</i>								
Projektgruppe	1.099	532.72 (2.915)	10.14**	.01	1.099	555.22 (3.134)	9.57**	.01
Vergleichsgruppe	871	517.78 (3.775)			870	539.64 (4.052)		
<i>Mathematikleistung (BB/BB)</i>								
Projektgruppe	389	492.38 (4.606)			389	512.55 (5.103)		
Vergleichsgruppe	383	447.22 (4.423)			383	464.50 (4.977)		
<i>Mathematikleistung (Hessen)</i>								
Projektgruppe	710	554.82 (3.472)	176.58***	.21	710	578.61 (3.682)	172.25***	.21
Vergleichsgruppe	488	573.15 (4.369)			487	598.72 (4.562)		
gesamt	1.970	526.11 (2.336)			1.969	548.34 (2.509)		

Anmerkungen. imputierter Datensatz, alle Werte gepoolt; Mittelwert (*MW*), Standardfehler des Mittelwertes (*SE*), F-Statistik aus ANOVA innerhalb des Messzeitpunktes mit Gruppenunterscheidung, η^2 als Maß der Effektstärke; Stichprobenumfang (*N*)

Diese Leistungsunterschiede werden grafisch besonders deutlich. Die Schülerinnen und Schüler der Bundesländer Berlin/Brandenburg befanden sich zum ersten Messzeitpunkt vollständig in der neunten Jahrgangsstufe. Ihr Mittelwert von $m=470$ Punkten befindet sich unterhalb des deutschen Durchschnittswertes von 500 Punkten. Die hessischen Schülerinnen und Schüler liegen mit einem Wert von 564 Punkten weit über dem deutschen Mittelwert. Das Wachstum der einzelnen Gruppen liegt zwischen 18 Punkten (Vergleichsschulen BE/BB) und 26 Punkten (Vergleichsschulen HE) und entspricht damit dem jährlich angenommen Kompetenzzuwachs von 20-30 Punkten.

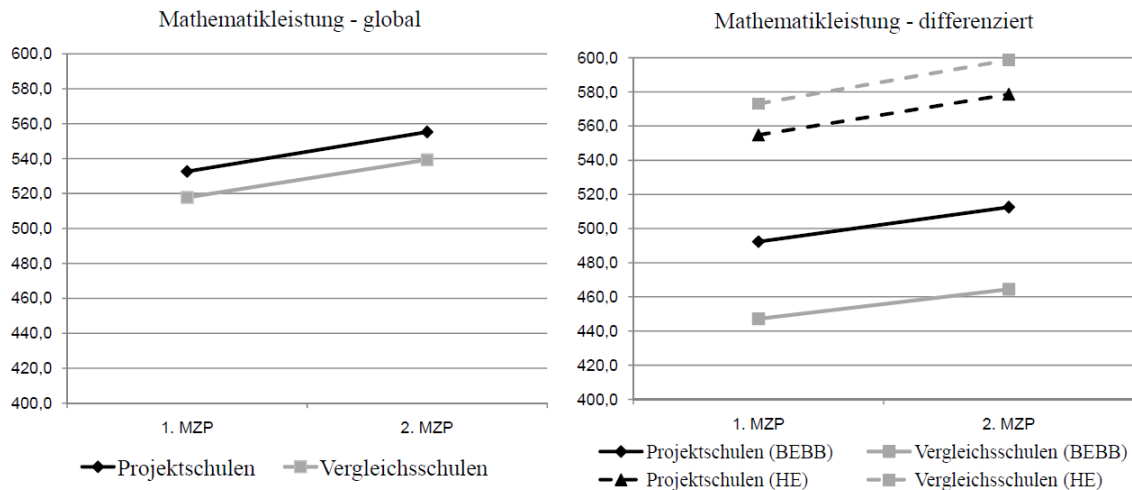


Abbildung 47. Veränderung in der Mathematikleistung, Globaldarstellung und detaillierte Darstellung.

Die Mehrebenenmodelle ermöglichen es jedoch, den unterschiedlichen Messzeitpunkten Rechnung zu tragen und diese in den Wachstumswerten zu berücksichtigen. Danach liegt die Mathematikleistung eines durchschnittlichen Schülers bei 453 Kompetenzpunkten und dabei macht es keinen Unterschied, ob er in einer Projekt- oder Vergleichsschule lernt. Allerdings hat das Bundesland einen großen Einfluss auf die Leistung eines Schülers. Im Bundesland Hessen liegen die Mathematikleistungen um 98 Punkte höher als in den Bundesländern Berlin und Brandenburg. Aus den Wachstumswerten wird ersichtlich, dass die Mathematikkompetenz durchschnittlich um 3 Kompetenzpunkte pro Monat wächst. Allerdings entwickeln sich die Kompetenzen der Schülerinnen und Schüler in den Projektschulen nicht signifikant schneller als in den Vergleichsschulen. Auch lernen die Schülerinnen und Schüler in Hessen nicht schneller als die Schüler in Berlin und Brandenburg. Signifikante Ergebnisse in den interessierenden Wachstumswerten sind für die Mathematikleistung nicht zu beobachten.

Tabelle 78. Wachstumseffekte in der Mathematikleistung der Schülerinnen und Schüler, Studie 2

		Parameter	Leistung
<i>Feste Effekte</i>			
Ausgangswert	Konstante	γ_{00}	453.195*** (4.529)
π_{0i}	Gruppe (0=VS, 1=PS)	γ_{01}	7.153 (4.793)
	Bundesland (0=BE/BB, 1=HE)	γ_{02}	98.138*** (4.974)
Wachstumswert	Konstante	γ_{10}	3.157*** (.419)
π_{1i}	Gruppe (0=VS, 1=PS)	γ_{11}	.045 (.404)
	Bundesland (0=BE/BB, 1=HE)	γ_{12}	-.243 (.449)
<i>Zufallseffekte</i>			
	innerhalb Personen	σ^2_{ε}	2002.450
Level 1	Ausgangswert	σ^2_0	7435.954***

Anmerkungen. VS = Vergleichsschulen, PS = Projektschulen, BE/BB = Berlin/Brandenburg, HE=Hessen, imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; Varianz des Wachstumsparameters auf Null gesetzt; ~p<.10; *p<.05; **p<.01; ***p<.001

6.3.3 Zusammenfassung der Forschungsfrage 2

Im Implementationsprozess der nationalen Bildungsstandards wurden zehn Merkmale an unterschiedlichen Stichproben untersucht, die für eine gelingende Umgestaltung des Unterrichts hin zu einer verstärkten Kompetenzorientierung als relevant angesehen werden. Die Forschungsfrage 1 nimmt die Wirkungen einer konzeptualisierten, unterstützenden Unterrichtsgestaltung auf die Lehrkräfte in den Blick und Forschungsfrage 2 fokussiert die Wahrnehmungen der Schülerinnen und Schüler, die sich einem, im Zuge der beschlossenen Bildungsstandards, veränderten Unterricht gegenübergestellt sehen. Im Vergleich der Wahrnehmungen der am Unterricht Beteiligten zeigt sich, dass die von den Lehrkräften intendierte Vorgehensweise von den Schülerinnen und Schülern nicht erkannt wird. Zu beachten bleibt, dass die Schülerinnen und Schüler der Studie 1 eine Negativauswahl im Hinblick auf ihre Mathematikleistungen gegenüber der Stichprobe im Ländervergleich darstellen. Vor diesem Hintergrund zeigen sich folgende Ergebnisse:

- Schülerinnen und Schüler mit schlechteren Mathematikleistungen, geringerem Kompetenzerleben, einem weniger stark ausgeprägten mathematischen Selbstkonzept und einem geringeren Sympathieempfinden gegenüber der Mathematiklehrkraft nehmen an der Studie nicht teil.

- Die Schülerinnen und Schüler der Projektschulen in Studie 1 nehmen zum Beginn der Studie einen signifikant weniger ausgeprägten kompetenzorientierten Unterricht in den Bereichen des Probleme lösens, modellierens, technischen arbeitens und kommunizierens.
- Die Schülerinnen und Schüler der Projektschulen in Studie 1 erleben eine Veränderung ihrer Mathematiklehrkraft, die sich signifikant negativ auf die Lehrerbewertung auswirkt.
- In den Projektschulen der Studie 1 nehmen die Schülerinnen und Schüler eine verstärkte Kompetenzorientierung im Bereich der Verwendung von mathematischen Darstellungen wahr.
- Studie 2 zeigt, dass die Wahrnehmungen der Schülerinnen und Schüler und der Lehrkräfte deutlich auseinander gehen.
- Für die Schülerinnen und Schüler in den Projektschulen in BB/BE führt das geänderte Unterrichtsangebot der Lehrkräfte nicht zu einer Steigerung der Lehrerwahrnehmung.
- Schülerinnen und Schüler der Projektschulen in BB/BE erleben einen weniger kompetenzorientierten Mathematikunterricht als zu Beginn. Die Kompetenzen des Argumentierens, Probleme Lösens und Modellierens werden weniger stark wahrgenommen.

7 Diskussion und Ausblick

Die Einführung einer top-down initiierten Bildungsinnovation vollzieht sich als Prozess, der je nach Umfang nicht selten einen Zeitraum von 10 Jahren in Anspruch nehmen kann. Auch bei der Implementation der nationalen Bildungsstandards war es unstrittig, dass eine Umgestaltung des Unterrichts auch mit erfahrenen Lehrkräften einen langen Zeitraum in Anspruch nehmen wird. Fraglich war nur, ob durch eine externe Unterrichtsentwicklungsmaßnahme der Prozess unterstützt werden kann, um die Abnehmer des Unterrichts für die Lösung mathematischer Problemstellungen zu befähigen. Das Projekt startete zu einem frühen Zeitpunkt im Implementationsprozess, zu dem die unterstützenden Angebote wie Lehrmaterialien, Bewertungsraster und Fortbildungen nur in begrenztem Umfang bereitstanden. Das Projekt zielte auf eine freiwillige Initiative der Projektschulen, einen frühen Schritt im ohnehin anstehenden Implementationsprozess zu machen. Die Befunde zur Teilnahmequote zeigen, dass das aus eigenem Antrieb entstehende Engagement und Interesse an einer Umgestaltung des Unterrichts nur knapp die Hälfte der Lehrerschaft trifft. Gründe für die fehlende Bereitschaft, sich mit einer Bildungsinnovation auseinanderzusetzen sind vielschichtig. So können strukturelle Beschlüsse der auf Systemebene zu einer Überfrachtung an Innovationsvorhaben führen. Aber auch organisatorische Zwänge auf Schulebene und die individuelle Einstellung der Lehrkräfte gegenüber der Innovation sind zu analysieren. Im Folgenden sollen die Ergebnisse der Studien kritisch diskutiert und in einen Gesamtzusammenhang gestellt werden.

7.1 Veränderung des Unterrichtsangebots

Zeitler et al. (2012) gingen im qualitativen Teilprojekt der Studie 1 der Frage nach, wie Lehrkräfte die Bildungsstandards verstehen und wie sie diese Innovation auf Schulebene annehmen. Für zwei der teilnehmenden Projektschulen in Studie 1 konnten die Forscherinnen eine heteronome Orientierung im Umgang mit den Bildungsstandards herausarbeiten. Es zeigte sich, dass die Lehrkräfte die Bildungsstandards nicht im Sinne eines kompetenzorientierten Unterrichts verstehen, sondern vielmehr als weitere Vorgabe der Bildungspolitik, die es einzuhalten gilt. „Das Neue wird dabei in den Kontext der alltäglichen Praxis und somit des Bekannten eingeordnet.“ (Zeitler et al., 2012, S. 131). Wenn sie in den Gruppendiskussionen erläutern, dass die Bildungsstandards als ein Instrument zur Vereinheitlichung des Unterrichts zu verstehen sind, dann ist hier das von Spillane et al. (2002) beschriebene „Sense-making“ (Kapitel 3.1.2) zu beobachten. Die Lehrkräfte

gleichen die zur Verfügung gestellten Informationen mit ihren bereits bestehenden Schemata ab und vollziehen eine Passung im Hinblick auf ihre Handlungsmöglichkeiten. Die Ergebnisse der Studie 1 zur Forschungsfrage 1 zeigen allerdings, dass die Lehrkräfte in den Projektschulen häufiger kompetenzorientierte Tätigkeiten initiieren, womit sich faktisch das Unterrichtsangebot veränderte. Ihre quantitativen Angaben geben jedoch keinen Aufschluss darüber, ob das Konzept der Bildungsstandards auch verstanden und didaktisch umgesetzt wurde. Vielmehr besteht für die Projektschulen der Studie 1 Anlass zu der Befürchtung, dass sich der Unterricht nicht verändert hat, sondern vielmehr nur kompetenzorientierte Aufgaben eingesetzt wurden. Die Schülerinnen und Schüler nehmen lediglich eine vermehrte Aufforderung wahr, mathematische Darstellungen zu verwenden. Diese Wahrnehmung könnte mit der Konstruktion kompetenzorientierter Aufgaben zusammenhängen, die sich aufgrund ihrer Praxisnähe oft bildlicher Darstellungen bedienen. Ein weiterführendes verändertes Unterrichtsangebot wird von den Schülerinnen und Schülern jedoch nicht wahrgenommen. Dieses Erkenntnis ist dahingehend plausibel, dass es sich bei den top-down eingeführten Bildungsstandards keinesfalls um ein fertiges Programm handelte, sondern vielmehr um eine Zielformulierung ohne Hinweise darauf, wie dieses Ziel zu erreichen sei. Es verlangt daher zunächst danach, Materialien zu entwickeln und zu erproben. Zum Zeitpunkt des Projekts lag als eines der wenigen Materialien das Buch „Bildungsstandards Mathematik: konkret“ von Blum et al. (2006) vor. Daher ist davon auszugehen, dass die Lehrkräfte in ihrem Unterricht vorerst nur die neuen Aufgaben sporadisch eingesetzt haben und sich noch in der Phase des Experimentierens und Ausprobierens befanden. Fullan (2007) führt aus, dass es wichtig ist, dass die Anwender die Innovation verstehen und über sie nachdenken (Abschnitt 3.1.1). Dieses Verständnis erfolgt bei den Lehrkräften der Projektschulen in Studie 1 nicht. Der Einsatz von kompetenzorientierten Aufgaben wird als etwas Zusätzliches oder Besonderes angeführt und stellt damit eine Abwechslung vom Unterrichtsalltag dar. Die anfangs starke Zunahme zum zweiten Messzeitpunkt und das teilweise Abfallen der Häufigkeit von kompetenzorientierten Tätigkeiten zum dritten Messzeitpunkt sind ein Hinweis darauf. Auch die gleichbleibend häufige Initiierung von Tätigkeiten des technischen Arbeitens deutet darauf hin, dass der Unterrichtsalltag in der Mehrheit beibehalten wurde. Natürlich bleibt zu beachten, dass die Skalen zum kompetenzorientierten Unterricht keine empirisch gebildeten Merkmale darstellen. Die Formulierungen ähneln im Wortlaut stark den Bildungsstandards und halten einer konfirmatorischen Faktorenanalyse nicht Stand. Teilweise be-

dingen sich die Kompetenzen gegenseitig und können daher nicht klar voneinander getrennt werden (Lenski et al., 2015). So stellt das Wissen um mathematische Routinen häufig die Grundlage für das mathematische Kommunizieren und Argumentieren dar. Da der Innovationsgehalt der Bildungsstandards für die Lehrkräfte nicht bedeutend ist („...der Innovationsgehalt der Bildungsstandards sei für dieses Kollegium lediglich die bewusste Fokussierung, die Inhalte würden „schon immer“ umgesetzt.“, Zeitler et al., 2012, S. 133), verfestigt und verstärkt sich auch ihre positive Einstellung gegenüber den Bildungsstandards nach 2-jähriger Projektlaufzeit (Studie 1), weil sie nur eine minimale Ergänzung dessen darstellen, was die Lehrkräfte in ihrem gesamten Berufsleben erfolgreich praktizieren. Unter Beachtung des zeitlichen Aspekts eines Implementationsprozesses könnte damit ein guter Ausgangspunkt geschaffen worden sein, um die eigentliche Umgestaltung des Unterrichts zu einer wirklichen Kompetenzorientierung anzugehen. Damit wird jedoch auch deutlich, wie zeitintensiv sich das Wirksamwerden einer Bildungsreform darstellt. Geht man davon aus, dass Lehrkräfte zunächst die neuen Aufgaben in ihrem Unterricht ausprobieren, dann zeigt bereits ein praxisnahes Rechenbeispiel den zeitlichen Umfang und die Grenzen einer wirklichen Unterrichtsveränderung: Zur Vermeidung von häufigen Lehrerwechseln in den Klassen unterrichten die Mathematiklehrkräfte eine Klasse meist über mehrere Jahre. Werden nun neue Aufgaben für eine bestimmte Zielgruppe (z.B. neunte Jahrgangsstufe, Stochastik) erprobt, dann umfasst ein günstiger Zeitraum bis zum zweiten, eventuell. angepassten Einsatz in einer neunten Klasse ein Schuljahr. Unterrichtet die Lehrkraft jedoch im nächsten Schuljahr nicht in einer neunten Klasse, sondern begleitet die Schülerinnen und Schüler vielmehr bis zum Ende der zehnten Klassenstufe und beginnt danach erneut mit einer siebten Klasse, umfasst der Zeitraum bis zum erneuten Einsatz der Aufgaben und die Beschäftigung mit den Potentialen der kompetenzorientierten Lernaufgaben drei weitere Schuljahre. Ein Wirksamwerden im intendierten Sinne der Reform ist im Kontext Schule folglich ein langfristiger Prozess, der von vielen Rahmenbedingungen abhängig ist (z.B. Einsatzplanung, Schülerklientel usw.). Rokitta et al. (2009) berichten aus der Schulpraxis, dass die Kollegien teilweise ganz unterschiedlich weit im Implementationsprozess vorangeschritten sind. Es fehlt an zeitlichen und fachlichen Ressourcen, um eine vertiefte Implementation zu erreichen. Als zeitlichen Druck erläutern die Schulleiter die Vielzahl von weiteren Innovationen in Bezug auf den Unterricht. Als fehlende fachliche Ressource werden die qualitativ höchst unterschiedlichen Fortbildungen zum Thema Bildungsstandards und Kompetenzorientie-

rung angeführt. Aus der wissenschaftlichen Begleitung der Studie 1 ist dieses Problem bekannt. Die Setkoordinatorinnen für die Projektschulen variierten das Konzept des IQB von einer Unterstützungsleistung der Eigenmotivation der Fachkonferenzen hin zur Durchführung von Fortbildungen. Aufgrund von Unsicherheiten bei den Lehrkräften im Umgang mit der Innovation entwickelten sich eher angeleitete Fortbildungen zum Thema. Zur Beantwortung der Frage, ob eine externe Unterstützung den Prozess der Implementation beschleunigen kann, sind weitere Aspekte heranzuziehen:

Die Einführung einer neuen Aufgabenkultur wird im SINUS-Programm als wesentlicher Gelingensfaktor im Unterrichtsentwicklungsprozess gesehen. Im Ergebnis erzielten die Schülerinnen und Schüler der beteiligten Schulen im Vergleich zu PISA-Schulen nicht nur höhere Kompetenzwerte in Mathematik und den Naturwissenschaften, sondern bewerteten ihren Unterricht auch qualitativ höher (Prenzel et al, 2005). Das Erfolgsrezept der SINUS-Schulen liegt in einer Kombination aus Kooperation innerhalb und zwischen den teilnehmenden Schulen, als auch in der Arbeit an didaktischen Modulen zur qualitativen Steigerung des Unterrichts im Hinblick auf den durch die Bildungsstandards gesetzten Referenzrahmen. Unterstützungssysteme spielen folglich bei der Veränderung einer routinierten Unterrichtspraxis eine entscheidende Rolle (Oelkers/Reusser, 2008, Ziener, 2006). Damit wird ein Moment in einem alignment-Prozess markiert, dessen Bezeichnung für einen Abstimmungsprozess zwischen accountability und Unterrichtsrealität steht (Berner/Stolz, 2006). Das *alignment* zwischen den verschiedenen Systemebenen und ihren Anforderungen an das Lehren und Lernen, die Abstimmung zwischen den Inhalten und Leistungen, auf die hin unterrichtet wird, und dem, was in Tests gemessen wird, kann in einem an Standards orientierten Bildungswesen als ein Maß für deren Implementation gelten. Maag Merki (2016) verweist auf die weitestgehend gesetzliche Verankerung der Bildungsstandards in Deutschland gegenüber Österreich und Schweiz. Mit der Gültigkeit der Bildungsstandards in allen Bundesländern, der Einführung eines systematischen Bildungsmonitoring auf der Grundlage der Bildungsstandards und zentralen Abschlussprüfungen scheint bereits ein hoher Grad im Implementationsprozess erreicht. Die Schülerleistungen in neun der 16 Bundesländer weisen im IQB-Ländervergleich aus dem Jahr 2012 Kompetenzwerte im Fach Mathematik über dem Durchschnittswert auf. Auch die Brandenburgischen Schülerinnen und Schüler erreichten im Ländervergleich einen Kompetenzwert von 518. Damit liegen sie um 18 Punkte über dem deutschen Durchschnitt und zeigen, dass sie kompetenzorientierten Aufgaben überdurchschnittlich gut begegnen

(Pant et al., 2013). Die Testergebnisse (accountability) weisen damit in eine positive Richtung. Es verbleibt ein Blick in die Unterrichtsrealität. Rokitta et al. (2009, S. 35) berichten vom unterschiedlichen Umgang mit den Bildungsstandards in der Weise, wie Zeitler et al. (2012) es ebenso in den verschiedenen Gruppendiskussionen rekonstruieren konnten: „Zum Einen wird die Idee der Kompetenzorientierung schon seit längerem [...] praktiziert und durch die Standards weiterentwickelt. Andere Kollegien sind „wachgerüttelt“ und arbeiten mit großem Elan an unterrichtlichen Veränderungen. Dritte wiederum formulieren bisheriges nur in Anstaltslehrplänen um und ändern nichts am Unterricht.“. Aber selbst wenn die Lehrkräfte in ihrem Verständnis der Bildungsstandards stark variieren, folgert Spillane (2004), muss die Weiterentwicklung des Unterrichts in Relation zum Ausgangsunterricht gesehen werden und deshalb stellt bereits die Verwendung anderer Aufgaben einen ersten Schritt der erfolgreichen Implementation dar. Folglich sind Angebote nach der Initiierung einer Bildungsreform, die die Gelegenheit zum Verständnis und einen Input mit Neuigkeitswert bieten als positiv zu bewerten. Die Ergebnisse der Evaluationsstudie (Studie 2) zeigen, dass die Schulen, in denen Entwicklungsprojekte laufen (Projektschulen), im Vergleich zu anderen Schulen in ihrer Region (Bundesland), den Bildungsstandards gegenüber positiver eingestellt sind, häufiger und auf höherem Niveau in der Fachkonferenz zusammenarbeiten und kooperieren und auch in der Unterrichtspraxis mehrheitlich einen stärker an den mathematischen Kompetenzen orientierten Unterricht anbieten.

Dennoch, die statistisch nachweisbaren Ergebnisse der Studien zeigen, dass es schwer ist, ein auf Freiwilligkeit basierendes Engagement zur Veränderung bei den Lehrkräften zu erzeugen, um die intendierten Ziele der Reform zu beschleunigen. Die erhofften Synergieeffekte im Kollegium blieben aus. Das zeigt der Vergleich der Daten der Lehrkräfte des Querschnitts und des Längsschnitts in Studie 1. Fraglich bleibt weiterhin, wie Lehrkräfte dazu motiviert werden können, bestehende Barrikaden aufzugeben und neue Konzepte anzugehen. Fullan (2007, S. 55) plädiert hierbei in Anlehnung an Elmore (2004a) für eine top-down-inszenierte Verordnung von Veränderung:

“The development of systematic knowledge about, and related to, large-scale instructional improvement requires a change in the prevailing culture of administration and teaching in schools. Cultures do not change by mandate; they change by the specific displacement of existing norms, structures, and processes by others; the process of cultural

change depends fundamentally on modeling the new values and behavior that you expect to displace the existing ones.” (Elmore, 2004a, S. 11).

Als Begründung führt Elmore (2004a) an, dass Lehrkräfte ein gegensätzliches Verständnis von Professionalität haben. Professionalität wird in der Schulpraxis daran festgemacht, wie autonom man agieren kann: inwiefern die Lernbedingungen selbst bestimmt werden können, inwiefern das Einzelurteil der Lehrkraft zur Leistungsbenotung entscheidend ist und die Unterrichtsatmosphäre geschützt ist.

Autonomie - als das Bedürfnis, sich als Urheber/in seines/ihres Verhaltens wahrzunehmen und den Interessen und integrierten Werten entsprechend handeln zu können (Deci & Ryan, 2008) – führt damit zu proaktivem Verhalten (Martinek, 2012). Die Einschränkung dieses pädagogischen Entscheidungs- und Handlungsspielraumes zählt nach Rudow (1994) zu den einflussreichsten Belastungsfaktoren für Lehrkräfte. Die selbstbestimmte Motivation zu unterrichten nimmt ab und sie verfallen in ein stark kontrolliertes Lehrverhalten, das sich wiederum negativ auf die Motivation der Schülerinnen und Schüler auswirkt (Pelletier et al., 2002). Martinek (2012) fand ein reziprokes Verhältnis zwischen der Befriedigung des Autonomiebedürfnisses von Lehrkräften und dem subjektiv empfundenen Druck im Lehrberuf. Die Forscherin diskutiert vor diesem Hintergrund auch die laufende Implementierung der Bildungsstandards in Österreich. Ähnlich wie in Deutschland werden Lehrpersonen bei der Entwicklung und Vermittlung der Standards miteinbezogen. Sie führt jedoch an, dass Partizipation und Wahlmöglichkeiten der Materialien nicht entscheidend für die Befriedigung des Autonomiebedürfnisses sind. Vielmehr sollten transparente Strukturen hinsichtlich der Ziele, der Abläufe und der geplanten Konsequenzen kommuniziert werden. Doch gerade die Ziele und Abläufe der seit Gültigkeit der Bildungsstandards von der Politik und Wissenschaft ausgerufenen Kompetenzorientierung sind unklar. Durch die Formulierung von Zielen bei weitestgehender Freiheit des Prozesses – ohne verabredete Konsequenzen – erscheint die Aufforderung zur Veränderung als Bitte. Daher zeigt sich auch im Antwortverhalten der Lehrkräfte der Studie 2, dass die Mehrheit der Lehrkräfte einen Verlust ihrer Autonomie nicht befürchtet. Lediglich 11 Prozent der Lehrkräfte aus BE/BB und 18 Prozent der Lehrkräfte aus Hessen sehen die Autonomie im Lehrberuf durch die Standards in Gefahr. Nach einem Schuljahr verschiebt sich die Verteilung auffällig für die Lehrkräfte in BE/BB. Mit zunehmendem Implementationsprozess ist eine Abkehr von der sorglos-Haltung zu beobachten. Die Verteilung verschiebt sich nach rechts (Abbildung 48). Dennoch verneinen mehr als die Hälfte der

befragten Lehrkräfte die Einschränkung der Handlungs- und Entscheidungsfreiheit in BE/BB. Die von Zeitler et al. (2012) entwickelte Typologie der Heteronomie scheint nicht generell übertragbar auf die Schulen in BE/BB. Allerdings könnte die weiterhin vorhandene sorglos-Haltung auch ein Beleg für das von Zeitler et al. (2012) rekonstruierte Verständnis sein, in dem die Lehrkräfte nur einen geringen Neuigkeitswert in der Reform sehen. Erst die Vielzahl von Berührungspunkten im Implementationsprozess bringt sie dazu, das Verständnis langsam zu verändern.

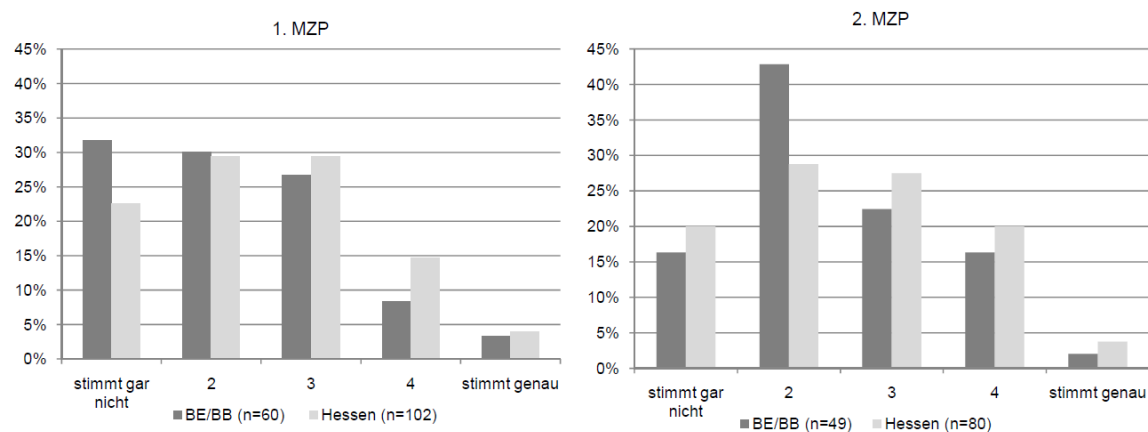


Abbildung 48. Antworthäufigkeiten in Prozent zum Item: "Die professionelle Handlungs- und Entscheidungsfreiheit wird durch Standards deutlich eingeschränkt.".

Das bestätigt sich auch bei einem Blick auf das Antwortverhalten zum Item: „Am ganzen Standard-Konzept ist noch Vieles ungeklärt, um es wirklich beurteilen zu können.“ Von einer anfänglichen Unbestimmtheit bei mehr als der Hälfte der Lehrkräfte in BE/BB verteilt sich das Antwortverhalten im 2. MZP nach links, wodurch mehr Klarheit und Verständnis geäußert wird (Abbildung 49). Die Lehrkräfte in BE/BB fühlen sich etwas sicherer im Umgang mit der Reform, wobei ein Großteil noch immer unentschlossen wirkt. Klare Handlungskonzepte könnten hier wirksam werden, wodurch die intendierte Gestaltungsfreiheit im Unterricht jedoch eingeschränkt wird. Zudem ist die durch den Ländervergleich 2012 ermittelte Altersstruktur in den Kollegien der neuen Bundesländer zu beachten. Es wurde ermittelt, dass gerade die Kollegien in den ostdeutschen Bundesländern mehrheitlich über 50 Jahre alt sind, was mit einer langjährigen didaktischen Erfahrung einhergeht. Somit ist zu vermuten, dass die Lehrkräfte ihre über Jahre eingeübten Unterrichtsroutinen nur schwerlich ändern. Weitere Unterrichtsreformen, wie die unter dem Oberbegriff „Inklusion“, überfrachten den Implementationsprozess und führen dazu, sich auf das bisher Bewährte zurückzubedenken, sofern keine Konsequenzen zu befürchten

sind. Mit den aktuellen Anpassungen der Lehrbücher ist zu vermuten, dass Lehrkräfte lediglich die neuen Aufgaben als Input benutzen, ohne eine intendierte Kompetenzorientierung herbeizuführen.

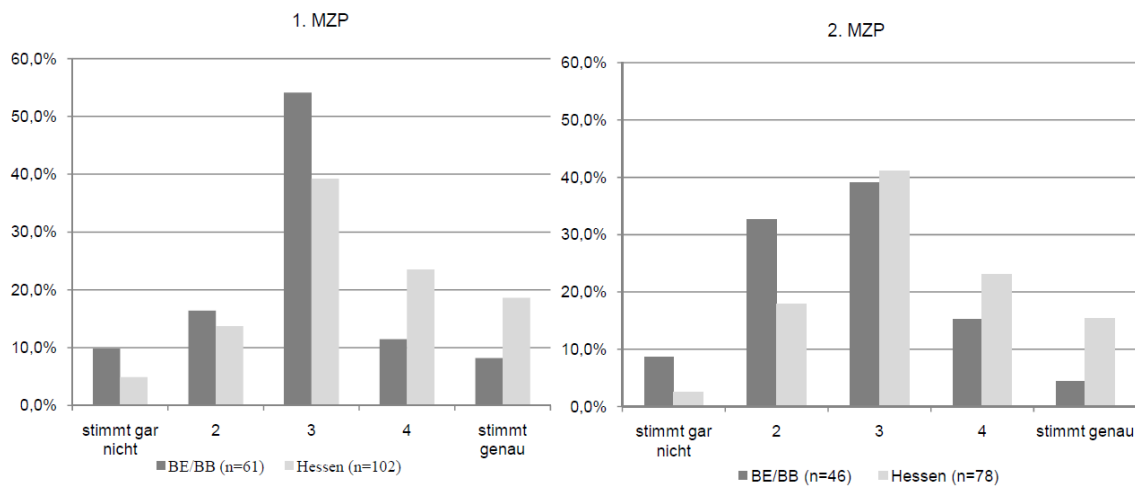


Abbildung 49. Antworthäufigkeiten in Prozent zum Item: "Am ganzen Standard-Konzept ist noch zu Vieles ungeklärt, um es wirklich beurteilen zu können.".

Grieshop (2009) beschreibt eine stärkere Kompetenzorientierung im Mathematikunterricht durch die Anwendung von vorgegebenen Instruktionen, die das Potential kompetenzorientierter Aufgaben ausschöpfen. Sie berichtet anhand von Beobachtungsbögen eine über das technische Arbeiten hinaus gehende Beschäftigung mit den intendierten Kompetenzbereichen im Gegensatz zum bloßen Rückgriff auf die kompetenzorientierten Schulbuchaufgaben. Es wird die Schlussfolgerung gezogen, dass Lehrkräfte zum Ausschöpfen des Potentials einer kompetenzorientierten Mathematikaufgabe Anleitung und Unterstützung benötigen. Die eigenen Ressourcen einer Lehrkraft enden beim bloßen Einsatz neuer Aufgaben, die jedoch weitestgehend in den Schulbüchern Einzug gehalten haben. Rokitta et al. (2009, S. 35) beschreiben dies als annähernden Abschluss der „Stufe der Bereitstellung von Gestaltungsmöglichkeiten“, woran sich die Gestaltungsfreiheit im Bereich der Unterrichtsprozesse durch die Lehrkräfte anschließen soll.

Gleichzeitig machen sie aber auch auf die Gefahr der Engführung des Unterrichts aufmerksam, der durch hierarchisierte Kompetenzraster ein lineares Lernen verlangt und andere (kreative) Kompetenzen vernachlässigt. Diese Gefahr bestätigt Maag-Merki (2016) für eine Stichprobe aus Bremen mit Blick auf das Zentralabitur. Die Lehrpersonen schränken in bedeutsam stärkeren Maß die Themenvarianz im Unterricht der zentral geprüften Kurse ein. Diese Schwerpunktlegung führte aber keinesfalls zu besseren Abitur-

leistungen, jedoch zu einer verringerten Lernmotivation der Schülerinnen und Schüler, da Alltagsbezüge und Interessen vernachlässigt wurden.

7.2 Veränderungen der Schülerwahrnehmung und der Mathematikleistung

Ein Vergleich der vorliegenden Befunde mit späteren Untersuchungen im Rahmen des Ländervergleichs 2012 zeigt, dass sich das mathematische Interesse der relevanten Schülerinnen und Schüler aus BE/BB und Hessen nicht verändert hat. Deutschland weist mit einem Bundesdurchschnitt von 2.51 auf einer vierstufigen Skala ein mittleres mathematisches Interesse auf. Die in dieser Arbeit ermittelten Werte liegen mit 2.03 (Studie 1, Projektgruppe, 3. MZP) und 2.23 (Studie 2, Vergleichsgruppe BE/BB, 1. MZP) unterhalb des Bundesdurchschnitts. Auch die Werte im Ländervergleich weisen sowohl für die Schülerinnen und Schüler aus BE/BB und Hessen ein unterdurchschnittliches mathematisches Interesse aus (Pant et al 2013). Hinsichtlich des Selbstkonzeptes wurde 2012 ein mittleres Selbstkonzept von 2.67 im Ländervergleich ermittelt. Die in den Studien 1 und 2 dieser Arbeit berechneten mittleren Werte im mathematischen Selbstkonzept liegen zwischen 2.39 (Studie 1, 3. MZP, imputierter Datensatz) und 2.55 (Studie 2, Projektgruppe Hessen, 2. MZP). Damit befinden sie sich unterhalb des Bundesdurchschnitts, aber dennoch im mittleren Niveau. Die Autoren des Ländervergleichs sehen keinen Widerspruch zwischen niedrigem Selbstkonzept und hohen mathematischen Leistungen bei den brandenburgischen Schülerinnen und Schülern. Als Erklärungsansatz dieses „scheinbar paradoxen Befundes“ (Jansen, Schroeders & Stanat, 2013, S. 361) werden Referenzrahmeneffekte angeführt, wonach die Selbsteinschätzung nicht nur nach den eigenen Leistungen erfolgt, sondern auch die Leistungen der Mitschüler und Gleichaltriger als Referenz hinzuzieht. Ebenso denkbar zur Erklärung der guten mathematischen Leistungen in Verbindung mit einem geringen mathematischen Selbstkonzept und Interesse wäre eine heteronome Arbeitshaltung – übertragbar von den Lehrkräften auf die Schülerinnen und Schüler. Durch die wiederholte Äußerung der in den Stoffverteilungsplänen vorgegebenen Kompetenzen durch die Lehrkräfte könnten die Schülerinnen und Schüler es als Ziel betrachten, die Vorgaben instruktionsartig umzusetzen, ohne ein echtes Verständnis und Interesse für die mathematischen Inhalte aufzubauen.

Die Schülerinnen und Schüler protokollieren auch keine Veränderung in der Unterrichtsgestaltung, allerdings sinken die Werte in der Lehrerbewertung in den Projektschulen im Verlauf der Studie 1. Während die Lehrerbewertung der Schülerinnen und Schüler in den Vergleichsschulen nahezu konstant auf gleichbleibendem Niveau verharrt, wird den

Lehrkräften in den Projektschulen eine negative Entwicklung attestiert. Folglich hat sich im Unterrichtsangebot doch etwas verändert, dem dadurch Rechnung getragen wird. Die Itemformulierungen der Skala zur Lehrerbewertung lassen darauf schließen, dass die Lehrkräfte die neue Unterrichtsgestaltung nur zögerlich ausprobieren oder negativer behaften (z.B. „Unser Mathematiklehrer/ unsere Mathematiklehrerin unterrichtet mit Begeisterung.“). Die Schülerinnen und Schüler könnten damit den Einsatz neuer, unangenehmer oder als schwierig empfundener Aufgaben bilanzieren. Zeitler et al. (2012) berichten von einer Gruppendiskussion in den Projektschulen, in der die Lehrerinnen den Unmut und die Ablehnung der Schülerinnen und Schüler gegenüber den anwendungsbezogenen Aufgaben beschreiben: „*Cf* greift das von *Af* geforderte praxisbezogene Denken der Schülerinnen und Schüler auf und nennt als Hinderungsgrund für den Einsatz von Aufgaben mit Anwendungsbezug im Unterricht, dass es „Widerstand“ gebe.“ (Zeitler et al, 2012, S. 135, Hervorhebungen im Original). Probleme mit den kompetenzorientierten Aufgaben zeigten sich auch bei den Abiturienten. Klein und Jahnke (2012) stellen beispielhaft dar, wie sich die Noten der Schülerinnen und Schülern verschieben, wenn sie eine kompetenzorientierte Mathematikaufgabe aus 2008 und eine Abituraufgabe aus dem gleichen Aufgabenbereich vor der Einführung des Zentralabitur von 2006 in NRW lösen. Es zeigte sich, dass die Noten bei der Aufgabe aus 2008 zwar positiver ausfielen, die kompetenzorientierte Aufgabe allerdings deutlich anspruchslosere Kenntnisse zum Lösen erforderte. „Bei den neuen Aufgabenstellungen scheitern allerdings viele Schüler an der ungewohnten Art der Fragestellung, die durch den Einsatz von Distraktoren gekennzeichnet ist.“ (Klein & Jahnke, 2012, S. 7). In Schülerkommentaren wird deutlich, dass in den ungenau formulierten Aufgabenstellungen häufig nach dem Sinngehalt gesucht wird. Außerdem vermuten die Schülerinnen und Schüler versteckte inhaltliche Schwierigkeiten in den meist textlastig präsentierten Aufgaben. Die Autoren schlussfolgern, dass durch kompetenzorientierte Aufgaben die Schülerleistungen im Fach Mathematik zwar verbessert werden, diese jedoch durch eine Nivellierung der Ansprüche erreicht wird. In der Wahrnehmung des Unterrichtsgeschehens werden den Schülerurteilen unterschiedliche Validitäten nachgesagt. Eine didaktische Bewertung fällt den Schülerinnen und Schülern zumeist schwer, da die von der Lehrkraft vorgenommenen Differenzierungen nicht erkannt werden. Schülerinnen und Schüler neigen zu einem Globalurteil über den Unterricht. Die Studien 1 und 2 stützen diesen bereits bekannten Befund, indem die Beurtei-

lungen in den Projektschulen sich nicht nennenswert von den Wahrnehmungen in den Vergleichsschulen unterscheiden.

Die aktuellste Vergleichsmöglichkeit hinsichtlich der Schülerleistungen im Fach Mathematik bietet der vom IQB durchgeführte Ländervergleich (Pant et al., 2013). Darin werden in den neuen Bundesländern Sachsen, Sachsen-Anhalt, Thüringen und Brandenburg und in Bayern Kompetenzwerte oberhalb des Bundesdurchschnitts referiert. Die Stadtstaaten weisen dagegen nur unterdurchschnittliche Kompetenzwerte auf. Weiterhin werden besonders in Berlin mit einem Prozentsatz von 38,5% Schülerinnen und Schüler identifiziert, die den Mindeststandard nicht erreichen. Für Brandenburg und drei weitere ostdeutsche Flächenbundesländer ist damit eine signifikante Steigerung gegenüber früheren Testungen (z.B. PISA 2006) zu verzeichnen. Die Schülerinnen und Schüler zeigen anhand ihrer Ergebnisse, dass sie kompetenzorientierte Aufgaben nunmehr überdurchschnittlich gut bearbeiten können. Wenngleich Drüke-Noe (2015) im Bereich des Modellierens darauf aufmerksam macht, dass die meisten Testaufgaben den Anforderungsbereichen I und II zuzuordnen sind, weisen die Ergebnisse dennoch darauf hin, dass die Schülerinnen und Schüler ihre mathematischen Kompetenzen verbessern bzw. ausbauen konnten. Unter der Annahme der als Input eingesetzten neuen Aufgaben – ohne intendierte Kompetenzorientierung – scheint diese Form der Implementation ebenso zielführend zu sein. In den vorliegenden Studien dieser Arbeit konnte eine Leistungssteigerung noch nicht gezeigt werden. Die Leistungswerte der Berliner und Brandenburger Neuntklässler liegen in Studie 1 um den Bundesdurchschnitt, in Studie 2 jedoch deutlich darunter. Nach weiteren vier bis fünf Jahren kann durch die Ergebnisse des Ländervergleichs 2012 gezeigt werden, dass durch ein verändertes Unterrichtsangebot (in Form von kompetenzorientierten Aufgaben) die Leistungen der Schülerinnen und Schüler verbessert werden können, wenngleich in Berlin und den anderen Stadtstaaten große zugewanderungsbedingte Disparitäten zu verzeichnen sind, die den mittleren Leistungswert senken.

Auch wenn es von den Vertretern aus Bildungsadministration, -politik und Hochschule gewünscht ist, die Kompetenzorientierung im Mathematikunterricht in intendierter Weise unter Ausschöpfung des Potentials kompetenzorientierter Lernaufgaben zu verändern, bleibt es doch eine Frage des Verständnisses und der Orientierung auf Seiten der Lehrkraft, inwieweit auf die Gegebenheiten der Einzelschule und Einzelstunde reagiert wird. Zeitler et al. (2012, S. 234) schlussfolgern ebenso: „..., dass die expliziten Bewertungen der Reform „Bildungsstandards“, die die Lehrkräfte in den Gruppendiskussionen äußern,

von den impliziten Orientierungen und Werthaltungen auf der habituellen Ebene [zu] unterscheiden [sind].“.

7.3 Grenzen bei der Messung von Unterrichtsmerkmalen

Die in diesen Studien vorgenommene Modellierung einer Kompetenzentwicklung verteilt einen fachthematisch abhängigen Kompetenzzuwachs, der in Studie 1 drei Mal erhoben und in Studie 2 zwei Mal gemessen wurde, anteilig auf einen definierten Zeitraum. Das bedeutet, der in den Modellen abgebildete monatliche Kompetenzzuwachs suggeriert ein stetiges Wachstum der abgebildeten Merkmale. Die in dieser Arbeit angewandte Methode der Längsschnittmodellierung mit HLM trägt bereits den unterschiedlichen Erhebungszeitpunkten Rechnung und vergrößert damit die Nähe zur Realität. Dennoch können die in der Wissenschaft angewandten Modelle nur versuchen, die Realität bestmöglich abzubilden. Gerade die Messung des Unterrichtsgeschehens stellt sich aufgrund der Vielzahl von Variablen als sehr komplex heraus. Die Ergebnisse können daher nicht allgemeingültig gelten. Eine andere Schulauswahl und eine andere Durchführung der Intervention in Form von begleitenden Mentoren würden das Ergebnis verändern. Die Datenlage machte es auch nicht möglich, die getesteten Klassen eindeutig einer Lehrkraft zuzuordnen. Um aussagekräftige Befunde zu erzielen, wäre dieser Umstand jedoch wünschenswert. Dann könnten Veränderungen in der Unterrichtsgestaltung direkt mit der Wahrnehmung einer bestimmten Schülergruppe in Verbindung gebracht werden. Ein solches Testdesign ist jedoch sehr kosten- und zeitintensiv. So wurden die Merkmale der Gruppe der Lehrkräfte und der Schülerinnen und Schüler hier getrennt ausgewertet.

Es ist nicht auszuschließen, dass die Lehrkräfte je nach Zuordnung zu Projekt- oder Vergleichsschulen in einer sozial erwünschten Form antworteten, wobei die Wahrscheinlichkeit bei den Projektschulen höher anzusetzen ist als bei den Vergleichsschulen, da diese ihre Gruppenzuordnung nicht kannten. So können im Abschlussbericht der Setkoordinatoreninnen Schulen ausgemacht werden, die mit Begeisterung im Projekt starteten und im Verlauf die Motivation aufgrund verschiedener Faktoren verringerten. Das Antwortverhalten der Lehrkräfte bei der Datenerhebung könnte einen mangelnden Umsetzungswillen der erarbeiteten kompetenzorientierten Lehr-Lernarrangements kompensieren. Die teilweise stark voneinander abweichenden Auskünfte der Lehrkräfte auf der einen Seite und der Schülerinnen und Schüler auf der anderen Seite geben einen Hinweis darauf. Dagegen spricht die von Clausen (2002) durchgeführte Studie, in der bereits nachgewiesen wurde, dass Schüler und Lehrkräfte nur sehr geringe Beurteilerübereinstimmungen aufweisen.

7.4 Ausblick

Die Ergebnisse weisen darauf hin, dass eine Veränderung des Unterrichts möglich ist. Es bleibt jedoch fraglich, ob die Lehrkräfte in den Projektschulen auch nach Ende des Projekts und der damit fehlenden externen Begleitung bzw. des Inputs ihren Unterricht weiterhin kompetenzorientiert gestalten. So wird im Abschlussbericht des LISUM darauf hingewiesen, dass der zeitliche Druck im Lehreralltag ein hinderlicher Faktor im Entwicklungsprozess war. Weitere Unsicherheiten wurden durch drohende Schulauflösungen und damit verbundene Umsetzungen, durch häufige Wechsel der Fachlehrer in der Einsatzplanung und die fehlende Unterstützung der Schulleitungen gesehen. Die Wahrscheinlichkeit einer weiterführenden Implementation vergrößert sich durch die konsequent durchgeführten Arbeiten im Bildungsmonitoring (Vergleichsarbeiten, Ländervergleiche) und die Reaktion der Schulbuchverlage auf die umgesetzten bundesweiten Beschlüsse der KMK. Insofern ist davon auszugehen, dass die Implementation der Bildungsstandards durch das Materialangebot der Verlage in Form von kompetenzorientierten Lernsituationen, Schulbüchern und Schülerarbeitsheften weiter voranschreitet. Der Grad der Implementation hängt jedoch von rein praktischen Rahmenbedingungen auf Schulebene ab. Bezogen auf die Fachschaft Mathematik bedeutet das konkret, dass bei unglücklicher Stunden- und Mitarbeiterplanung eine Verstetigung des kompetenzorientierten Unterrichts sehr lang dauern kann. So wird die Lehrkraft im besten Fall jährlich für den Unterricht in einer bestimmten Klassenstufe eingeplant und kann ihre im letzten Jahr ausgetestete kompetenzorientierte Lerneinheit weiter optimieren. Die Innovation Bildungsstandards verliert in diesem Szenario bereits nach einigen Jahren ihren Status. In der Praxis wird ein häufiger Lehrerwechsel jedoch vermieden, wodurch eine Lehrkraft eine bestimmte Klasse über mehrere Jahre begleitet. Das bedeutet jedoch, dass eine zur Erprobung konzipierte Lerneinheit für die achte Klassenstufe im schlechtesten Fall erst wieder in fünf Jahren weiterentwickelt werden kann, dann nämlich, wenn die Lehrkraft die Klasse zum Abschluss der 10. Jahrgangsstufe geführt hat und mit der neuen fünften Klasse (in Berlin und Brandenburg siebte Klasse) wieder inhaltlich an dem Punkt des Curriculums ankommt. Folglich sind die Ergebnisse des Projektes auch als solche zu verstehen. Tragfähige Aussagen über den Grad der Kompetenzorientierung im Mathematikunterricht würden kostenintensive Längsschnittstudien in Kombination mit qualitativ ausgewerteten Interviews und Stundenaufzeichnungen liefern.

Aus ökonomischer Sicht könnte die Definition einer Zielgruppe zur Beschleunigung einer Unterrichtsveränderung beitragen. In der Praxis orientieren sich junge Lehrkräfte in den

ersten Jahren stark an den erfahrenen Kollegen. Obwohl die Lehrerausbildung bereits gut auf die Kompetenzorientierung in den Klassenzimmern hin ausbildet, erfolgt in den ersten Arbeitsjahren als Lehrkraft eine Integration in das Kollegium der Einsatzschule. Es gilt, Material zu sammeln und die Unterrichtsstunden zumeist sinnvoll zu füllen. Dabei wird in erster Linie darauf hingearbeitet, die eigene Rolle im Umgang mit den Schülerinnen und Schülern zu definieren. Je nach Stundendeputat und entsprechend zur Verfügung stehenden Zeit liegen die Prioritäten bei der Aneignung der fachlichen Erfordernisse. Die relevante Zielgruppe der 30- bis 50-jährigen hat ihre Lehrerpersönlichkeit bereits definiert und genügend Material zur Durchführung eines gehaltvollen Unterrichts vorbereitet. Oftmals verlagern sich in dieser Altersgruppe jedoch die Prioritäten ins Private. Die Zeit, die vorhandenen Materialien auf eine höhere Kompetenzorientierung umzustellen, fehlt genauso wie eine subjektiv empfundene Notwendigkeit. In dieser Altersgruppe liegt jedoch das größte Potential, eine Innovation zur Verstetigung voranzutreiben. Hier müssten Kapazitäten geschaffen werden, die es ermöglichen, einen empfundenen Mehraufwand in die übliche Arbeitszeit zu integrieren. Älteren Kollegen erscheint eine solche Arbeitsinvestition als nicht ertragsbringend im Verhältnis zu ihren verbleibenden Arbeitsjahren. Der Einsatzplanung in den Schulen kommt damit eine hohe Bedeutung zu.

Empfehlenswert für den seit nunmehr zehn Jahre andauernden Implementationsprozess der Bildungsstandards scheint eine stärkere Steuerung oder eine neue Aktualität. Die Studie von Richter et al. (2014) zeigt, dass Lehrkräfte in den Grundschulen in Abhängigkeit ihrer Überzeugungen zu den Vergleichsarbeiten den Unterricht nicht verändern. Werden die Vergleichsarbeiten von den Lehrkräften als Mittel der Unterrichtsentwicklung begriffen, erfolgt eher eine verstärkte Kompetenzorientierung im Unterricht als bei der Wahrnehmung der Vergleichsarbeiten als Kontrollinstrument. Je nach Überzeugung der Lehrkräfte fanden die Autoren Hinweise auf eine Verengung des Curriculums. Eine solche Fokussierung auf die in Tests erfragten Kompetenzen birgt die Gefahr, notwendige Grundfertigkeiten zu vernachlässigen. Der intendierte Aufbau von Modellierungs- und Problemlösekompetenzen im Mathematikunterricht könnte in einem Widerspruch dazu stehen, die Kompetenzen im technischen Arbeiten nicht weiter auszubauen. Doch gerade das wiederholte Einüben von Rechenroutinen verschafft die nötige Basis zum Verständnis übergeordneter Modelle und Argumentationen. Auch das abnehmende System der Berufsausbildung ist auf die ausreichende Ausbildung technischer Arbeitsroutinen angewiesen. Gerade im Bereich der Berufsausbildung sind elementare mathematische Routinen

wie die Dreisatz- und Prozentrechnung essentiell für den Erfolg. Eine in der Berufsausbildung zweifelsohne von den Auszubildenden erkannte Relevanz der Aufgaben, führt ohne Kompetenzen im technischen Arbeiten leicht zur Demotivation. Die Stimmen der Industrie werden lauter, dass es nicht genügend geeignete Bewerber für den Fachkräftenachwuchs gibt. Häufig scheitert ein Ausbildungsvertrag an unzureichenden Mathematikkenntnissen. Für das allgemeinbildende Schulsystem bedarf es möglicherweise einer erneuten Aufmerksamkeit auf die Kompetenzorientierung, um die Lehrkräfte für die intendierte Wirkung der Bildungsstandards erneut zu sensibilisieren, ohne die parallel konzipierten Konzepte zum Umgang mit Flüchtlingen bzw. zur Inklusion zu verdrängen.

8 Literatur

- Altrichter, H. (2008). Veränderungen der Systemsteuerung im Schulwesen durch die Implementation einer Politik der Bildungsstandards. In T. Brüsemeister & K.-D. Eubel (Hrsg.), *Evaluation, Wissen und Nichtwissen* (S. 75-115). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Altrichter, H. (2010). Lehrerfortbildung im Kontext von Veränderungen im Schulwesen. In F. H. Müller, A. Eichenberger, M. Lüders, & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung* (S. 17-34). Münster: Waxmann.
- Altrichter, H. & Maag Merki, K. (Hrsg.). (2010). *Handbuch Neue Steuerung im Schulsystem*. Wiesbaden: Verlag für Sozialwissenschaften.
- Altrichter, H. & Posch, P. (2007). Analyse erster Erfahrungen mit der Implementation von Bildungsstandards. *Erziehung und Unterricht*, 157, S. 654-671.
- Artelt, C. & Riecke-Baulecke, T. (2004). Bildungsstandards. München: Oldenbourg.
- Asbrand, B., Heller, N. & Zeitler, S. (2012). Die Arbeit mit Bildungsstandards in Fachkonferenzen. Ergebnisse aus der Evaluation des KMK-Projektes *for.mat*. *Die deutsche Schule*, 104/1, S. 31-42.
- Avenarius, H., Ditton, H., Döbert, H., Klemm, K., Klieme, E., Rürup, M., Tenort, H.-E., Weishaupt, H. & Weiss, M. (2003). *Bildungsbericht für Deutschland. Erste Befunde*. Opladen: Leske+Budrich.
- Baumert, J. & Köller, O. (2000). Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III – Dritte Internationale Mathematik und Naturwissenschaftsstudie. Band 2* (S. 271-315). Opladen: Leske + Budrich.
- Baumert, J. & Kunter, M. (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9. Jg., S. 469-520.

- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.) (2000). *TIMSS/III – Dritte Internationale Mathematik und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn; Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit; Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe*. Opladen: Leske + Budrich.
- Baumert, J., Brunner, M., Lüdtke, O. & Trautwein, U. (2007). Was messen internationale Schulleistungsstudien? – Resultate kumulativer Wissenserwerbsprozesse. Eine Antwort auf Heiner Rindermann. *Psychologische Rundschau*, 58 (2), S. 118-128.
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-J. Rolff, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 314-354). Münster: Waxmann.
- Baumert, J., Lüdtke, O., Trautwein, U. & Brunner, M., (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review*, 4, 165-176.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten: Institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert et al. (Hrsg.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 259-330). Opladen: Leske+Budrich.
- Beer, R. (2007). *Bildungsstandards. Einstellungen von Lehrerinnen und Lehrern*. Münster: Lit.
- Berkemeyer, N., Bos, W., Manitius, V. & Müthing, K. (Hrsg.) (2008), *Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven*. Münster: Waxmann.
- Berner, E. & Stolz, S. (2006). *Literaturanalyse zu Entwicklung, Anwendung und insbesondere Implementation von Standards in Schulsystemen: Nordamerika*. Im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK. Zürich.

- Berner, E., Oelkers, J. & Reusser, K. (2008). Implementation von Bildungsstandards: Bedingungen des Gelingens (und Scheiterns) aus internationaler Sicht. *Zeitschrift für Pädagogik*, 53. Beiheft, S. 210-226
- Bethge, T. & Priebe, B. (2009). Unterrichtsentwicklung und Schulentwicklung. In: U. Klinger (Hrsg.), *Mit Kompetenz Unterricht entwickeln; Fortbildungskonzepte und –materialien* (S 69-100). Troisdorf: Bildungsverlag Eins.
- Biehler, R & Hartung, R. (2006). Die Leitidee Daten und Zufall. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 51-80). Berlin: Cornelsen Scriptor.
- Bijleveld, C. J. H. & van der Kamp, L. J. T. (1998). *Longitudinal Data Analysis. Designs, Models and Methods*. London: Sage.
- Bitan-Friedlander, N., Dreyfus, A. & Milgrom, Z. (2004). Types of „teachers in training“: the reactions of primary school science teachers when confronted with the task of implementing an innovation. *Teaching and Teacher Education*, 20, S. 607–619.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S.W. Kozlowski (Hrsg.), *Multilevel theory, research, and methods in organizations* (S. 349–381). San Francisco, CA: Jossey-Bass.
- BLK (1997). *Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“*. Heft 60 der Materialien zur Bildungsplanung und Forschungsförderung. Bonn: Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung.
- Blömeke, S., Eichler, D. & Müller, C. (2003). Rekonstruktion kognitiver Strukturen von Lehrpersonen als Herausforderung für die empirische Unterrichtsforschung. Theoretische und methodologische Überlegungen zu Chancen und Grenzen von Videostudien. *Unterrichtswissenschaft*, 31, S. 103-121.
- Blum, W. (2006a). Einführung. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 14-32). Berlin: Cornelsen Scriptor.

- Blum, W. (2006b). Die Bildungsstandards Mathematik. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 14-32). Berlin: Cornelsen.
- Blum, W., Drücke-Noe, C., Hartung, R. & Köller, O. (Hrsg.). (2006). *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen*. Berlin: Cornelsen.
- Böttcher, W., Bos, W., Döbert, H., & Holtappels, H. G. (Hrsg.). (2008). *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*. Münster: Waxmann.
- Böttcher, W. & Dicke, J. N. (2008). Implementation von Standards. In W. Böttcher, W. Bos, H. Döbert & H. G. Holtappels (Hrsg.), *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive* (S. 143-156). Münster: Waxmann.
- Bonsen, M. (1995). Lerngemeinschaften in der Schule. In H.G. Holtappels & K. Höhmann (Hrsg.), *Schulentwicklung und Schulwirksamkeit* (S. 180-195). Weinheim: Juventa.
- Bonsen, M. & Berkemeyer, N. (2011). Lehrerinnen und Lehrer in Schulentwicklungsprozessen. In E. Terhart, H. Bennewitz, & M. Rothland. (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 731-747). Münster: Waxmann.
- Bonsen, M. & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. *Zeitschrift für Pädagogik*, 52 (2), S. 167-184.
- Bos, W., Klieme, E. & Köller, O. (Hrsg.) *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert*. Münster: Waxmann.
- Bremerich-Vos, A., Böhme, K. & Robitzsch, A. (2009). Sprachliche Kompetenzen im Fach Deutsch - Strukturanalysen und Validierungsbefunde. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther. *Bildungsstandards in Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 198-218) Weinheim: Beltz.
- Bruder, R. (2000). Mit Aufgaben arbeiten. *Mathematik lehren*, 101, S. 12-17.
- Bruder, R. (2006). Langfristiger Kompetenzaufbau. In W. Blum, C. Drücke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret* (S. 135-151). Berlin: Cornelsen Scriptor.

- Bruder, R., Büchter, A. & Leuders, T. (2008). *Mathematikunterricht entwickeln. Bausteine für kompetenzorientiertes Unterrichten*. Berlin: Cornelsen Scriptor.
- Brüsemeister, T. & Eubel, K.-D. (Hrsg.) (2008), *Evaluation, Wissen und Nichtwissen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Brunner, M. (2006). Mathematische Schülerleistung: Struktur, Schulformunterschiede und Validität. Dissertation, Humboldt-Universität zu Berlin. Verfügbar unter <http://edoc.hu-berlin.de/dissertationen/brunner-martin-2006-02-08/PDF/brunner.pdf>.
- Brunner, M. & Krauss, S. (2010). Modellierung kognitiver Kompetenzen von Schülern und Lehrkräften mit dem Nested-Faktormodell. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 105-126). Münster: Waxmann.
- Brunner, M., Kunter, M., Krauss, S., Klusmann, U., Baumert, J., Blum, W. et al. (2006). Die professionelle Kompetenz von Mathematiklehrkräften: Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. Eine Zwischenbilanz des COACTIV-Projekts. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schulen. Abschlussbericht des DFG- Schwerpunktprogramms* (S. 54–82). Münster: Waxmann.
- Bryk, A., Camburn, E., & Louis, K. S. (1999). Promoting school improvement through professional communities: An analysis of Chicago elementary schools. *Educational Administration Quarterly*, 35(4), 707-750.
- Büchter, A. & Leuders, T. (2005). *Mathematikaufgaben selbst entwickeln. Lernen fördern – Leistung überprüfen*. Berlin: Cornelsen Scriptor.
- Chohors-Fresenborg, E., Sjuts, J. & Sommer, N. (2004). Komplexität von Denkvorgängen und Formalisierung von Wissen. In M. Neubrand (Hrsg.), *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000* (S. 109-144). Wiesbaden: Verlag für Sozialwissenschaften.
- Chomsky, N. (1973). *Aspekte der Syntax-Theorie*. Frankfurt/M.: Suhrkamp
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Münster: Waxmann.

- Cohen, D.K. and Barnes, C.A. (1993a). Conclusion: A new pedagogy for policy? In D. K. Cohen, M. W. McLaughlin & J. E. Taulbert (Hrsg.), *Teaching for Understanding: Challenges for Policy and Practice* (S. 240-275). San Francisco: Jossey-Bass Inc.
- Cohen, D. K., McLaughlin, M. W. & Taulbert, J. E. (Hrsg.) (1993), *Teaching for Understanding: Challenges for Policy and Practice* . San Francisco: Jossey-Bass Inc.
- Deci, E.L. & Ryan, R.M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology*, 49 (3), S. 182-185.
- Demuth, R., Gräsel, C., Ralle, B. & Parchmann, I. (Hrsg.) (2008), *Chemie im Kontext - Von der Innovation zur nachhaltigen Verbreitung einer Unterrichtskonzeption*. Münster: Waxmann.
- Desimone, L. (2002). How Can Comprehensive School Reforms Models Be Successfully Implemented? *Review of Educational Research*, 72 (3), S. 433-479.
- Desimone, L., Porter, A. C., Garet, M., Suk Yoon, K. & Birman, B. (2002): Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis*, 24, (2), S. 81-112.
- Diedrich, M., Thußbas, C. & Klieme, E. (2002). Professionelles Lehrerwissen und selbstberichtete Unterrichtspraxis im Fach Mathematik. *Zeitschrift für Pädagogik*, 45. Beiheft, S. 107-123.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der Forschung. In A. Helmke, W. Hornstein, E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich: Schule Sozialpädagogik, Hochschule*. 41. Beiheft der Zeitschrift für Pädagogik (S. 73-92). Weinheim und Basel: Beltz.
- Drieschner, E. (2009). *Bildungsstandards praktisch. Perspektiven kompetenzorientierten Lehrens und Lernens*. Wiesbaden: Verlag für Sozialwissenschaften.
- Drücke-Noe, C. (2015). Bildungsstandards und Modellieren: Wo stehen wir? In G. Kaiser, H.-W. Henn (Hrsg.), *Werner Blum und seine Beiträge zum Modellieren im Mathematikunterricht. Festschrift zum 70. Geburtstag von Werner Blum* (S. 89-104). Wiesbaden: Springer Spectrum.
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. *Zeitschrift für Pädagogische Psychologie*, 22. Jg, S. 193-206.

- DuFour, R. & Eaker, R. (1998). *Professional Learning Communities at Work. Best Practices for Enhancing Student Achievement*. Bloomington: National Education Service.
- Ehmke, T., Leiß, D., Blum, W. & Prenzel, M. (2006). Entwicklung von Testverfahren für die Bildungsstandards Mathematik. Rahmenkonzeption, Aufgabenentwicklung, Feld- und Haupttest. *Unterrichtswissenschaft*, 34 (3), S. 220-238.
- Elmore, R. (2004a). *The hollow core of leadership in practice*. Unpublished Paper. Cambridge, MA: Harvard University Graduate School of Education.
- Euler, D. & Sloane, T.F.E. (1998). Implementation als Problem der Modellversuchsforschung. *Unterrichtswissenschaft*, 26(4), S. 312 – 326.
- Faulstich-Christ, K., Lersch, R. & Moegling, K. (Hrsg.) (2010). *Kompetenzorientierung in Theorie, Forschung und Praxis*. Immenhausen bei Kassel.
- Fend, H. (1981). *Theorie der Schule*. München: Urban & Schwarzenberg.
- Forgasz, H. & Leder, G.C. (2008). Beliefs about mathematics and mathematics teaching. In P. Sullivan & T. Wood (Hrsg.), *International handbook of mathematics teacher education. Vol.1: Knowledge and beliefs in mathematics teaching and teaching development* (S.173-192). Rotterdam: Sense Publishers.
- Freudenthal, H. (1977). *Weeding and Sowing. Preface to a Science of Mathematical Education*. Dordrecht: D. Reidel Publishing Company.
- Freudenthaler, H. H. & Specht, W. (2006): *Bildungsstandards: Der Implementationsprozess aus der Sicht der Praxis. Ergebnisse einer Fragebogen-Studie nach dem ersten Jahr der Pilotphase II*. Graz: Zentrum für Schulentwicklung. Bundesministerium für Bildung, Wissenschaft und Kultur.
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., Pekrun, R. (2009). *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Fullan, M. (1982). *The meaning of educational change*. New York: Teachers College Press.
- Fullan, M. (1993). *Change forces: Probing the depths of educational reform*. New York: Falmer Press.
- Fullan, M. (1994). Coordinating top-down and bottom-up strategies for educational reform. In R. F. Elmore & S. H. Fuhrman (Hrsg.), *The governance of curriculum* (S.

- 186-202). Alexandria, VA: Association for Supervision and Curriculum Development.
- Fullan, M. (2001). *Leading in a Culture of Change*. San Francisco: Jossey-Bass.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York: Teachers College Press.
- Fussangel, K. & Gräsel, C. (2008). Unterrichtsentwicklung in Lerngemeinschaften: das Beispiel „Chemie im Kontext“. In N. Berkemeyer, W. Bos, V. Manitius & K. Müthing (Hrsg.), *Unterrichtsentwicklung in Netzwerken. Konzeptionen, Befunde, Perspektiven* (S. 285-295). Münster: Waxmann.
- Fussangel, K., Rürup, M. & Gräsel, C. (2010). Lehrerfortbildung als Unterstützungssystem. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 327-354). Wiesbaden: Verlag für Sozialwissenschaften.
- Gärtner, H. (2007). *Unterrichtsmonitoring*. Münster: Waxmann.
- Gillen, J. & Kaufhold, M. (2005). Kompetenzanalysen – kritische Reflexion von Begrifflichkeiten und Messmöglichkeiten. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 101 (3), S. 364-378.
- Graham, J. W. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, 60, S. 549-576.
- Granzer, D., Köller, O., Bremerich-Vos, A., van den Heuvel-Panhuizen, M., Reiss, K. & Walther, G. (Hrsg.). *Bildungsstandards in Deutsch und Mathematik. Leistungsmessung in der Grundschule*. Weinheim: Beltz.
- Gräsel, C. (2010). Stichwort: Transfer und Transferforschung im Bildungsbereich. *Zeitschrift für Erziehungswissenschaft*, 13, S. 7-20.
- Gräsel, C. & Parchmann, I. (2004). Implementationsforschung – oder: der steinige Weg, Unterricht zu verändern. *Unterrichtswissenschaft*, 32 (3), S. 196-214.
- Gräsel, C., Pröbstel, C., Freienberg, J. & Parchmann, I. (2006). Anregung zur Kooperation von Lehrkräften im Rahmen von Fortbildung. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 310-332). Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann.
- Grieshop, G. (2009). *Das Projekt „Schulbuch KO“ – Schulbuchaufgaben kompetenzorientiert einsetzen*. Beiträge zum Mathematikunterricht 2009 Online. Vorträge auf der 43. Tagung für Didaktik der Mathematik. Jahrestagung der Gesellschaft für Didaktik der Mathematik vom 02.03. bis 06.03.2008 in Oldenburg. (2009).

- Grillitsch, M. BIFIE-Report 6/2010. *Bildungsstandards auf dem Weg in die Praxis. Ergebnisse einer Befragung von Lehrkräften und Schulleiter/innen der Sekundarstufe I zur Rezeption der Bildungsstandards und deren Implementation*. Graz: Leykam.
- Groß Ophoff, J., Hosenfeld, I. & Koch, U. (2006). Ergebnissrückmeldungen und ihre Rezeption im Projekt VERA. In H. Kuper & J. Schneewind (Hrsg.), *Rückmeldung und Rezeption von Forschungsergebnissen – Zur Verwendung wissenschaftlichen Wissens im Bildungssystem* (S. 19-40). Münster: Waxmann.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Hall, G.E. & Hord, S.M. (2006). *Implementing Change. Patterns, Principles and Pot-holes*. Boston: Pearson.
- Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., Klein, S. P., Robyn, A. & Bugliari, D. (2003). Studying Large-Scale Reforms of Instructional Practice: An Example from Mathematics and Science. *Educational Evaluation and Policy Analysis*, 25 (1), S. 1-29.
- Hanson M. R., Hayes, j. R., Schriver, K., LeMahieu, P. G. & Brown, P. J. (1998). *A Plain Language Approach to the Revision of Test Items*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Hartig, J. & Jude, N. (2007). Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik* (S. 17-36). Bundesministerium für Bildung und Forschung, Band 20. Bonn, Berlin.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Heidelberg: Springer.
- Hartig, J. & Klieme, E. (Hrsg.) (2007). *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik*. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung. Band 20. Bonn.
- Hartig, J. & Kühnbach, O. (2006). Schätzung von Veränderung mit Plausible Values in mehrdimensionalen Rasch-Modellen. In: A. Ittel & H. Merkens (Hrsg.), *Veränderungsmessung und Längsschnittstudien in der Erziehungswissenschaft* (S. 27-44). Wiesbaden: Verlag für Sozialwissenschaften.

- Hasselhorn, M. & Gold, A. (2006). *Pädagogische Psychologie. Erfolgreiches Lernen und Lehren*. Stuttgart: Kohlhammer.
- Hattie, J. A. C. (2009). *Visible Learning. A synthesis of over 800 meta-analyses relating to achievement*. London und New York: Routledge.
- Hattie, J. A. C., Beywl, W. & Zierer, K. (2013). *Lernen sichtbar machen*. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“. Baltmannsweiler: Schneider Verlag Hohengehren.
- Helmke, A. (1999). Direkte Instruktion – effektiver Unterricht? *Bildung Real*, 38, 59-72.
- Helmke, A. (2004). Von der Evaluation zur Innovation: Pädagogische Nutzbarmachung von Vergleichsarbeiten in der Grundschule. *Seminar*, 2, S. 90-112.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.
- Helmke, A. & Hosenfeld, I. (2004). Vergleichsarbeiten – Standards – Kompetenzstufen: Begriffliche Klärung und Perspektiven. In R. S. Jäger, A. Frey & M. Wosnitza (Hrsg.), *Lernprozesse, Lernumgebungen und Lerndiagnostik. Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert* (S. 56-75). Landau: Verlag Empirische Pädagogik.
- Helmke, A., Hosenfeld, I., Schrader, F.-W. & Wagner, W. (2002). Unterricht aus der Sicht der Beteiligten. In A. Helmke & R. S. Jäger (Hrsg.), *Das Projekt MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext* (S.325-411). Landau: Verlag Empirische Pädagogik.
- Helmke, A. & Jäger, R. S. (Hrsg.) (2002). *Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext*. Landau: Verlag Empirische Pädagogik.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich D Praxisgebiete, Serie I Pädagogische Psychologie, Band 3* (S. 71-176). Göttingen: Hogrefe.
- Herzog, W. (2010). Besserer Unterricht dank Bildungsstandards und Kompetenzmodellen? In A. Gehrman, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Eine Verbesserung der Qualität von Schule, Unterricht und Lehrerbildung* (S. 37-46). Bad Heilbrunn: Klinkhardt.

- Hessisches Kultusministerium (Hrsg.) (2008). *Fortbildungshandreichung zu den Bildungsstandards Mathematik*. Sekundarstufe I. Wiesbaden: Hessisches Kultusministerium.
- Hochuli, L. A. (2011). *Bildungsstandards. Der Blick von Experten auf die Entwicklung von Bildungsstandards in der Schweiz und die Folgen für die Akteurkonstellation im Schulsystem*. Dissertation Universität Giessen.
- Holtappels, H. G. (2005). Bildungsqualität und Schulentwicklung. In H. G. Holtappels & K. Höhmann (Hrsg.), *Schulentwicklung und Schulwirksamkeit* (S. 27-47). München: Weinheim.
- Holtappels, H. G. & Höhmann, K. (Hrsg.), *Schulentwicklung und Schulwirksamkeit*. Weinheim, München: Juventa.
- Hord, S. (1997). *Professional Learning Communities: Communities of Continuous Inquiry and Improvement*. Southwest Educational Development Laboratory. Austin, Texas.
- Hosenfeld I. & Schrader, F. W. (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven*. Münster: Waxmann.
- Hugener, I., Pauli, C. & Reusser, K. (2007). Inszenierungsmuster, kognitive Aktivierung und Leistung im Mathematikunterricht. Analysen aus der schweizerisch-deutschen Videostudie. In D. Lemmermöhle, M. Rothgangel, S. Bögeholz, M. Hasselhorn & R. Watermann (Hrsg.), *Professionell lehren, erfolgreich lernen* (S. 109-122). Münster: Waxmann.
- Jablonka, E & Keitel, C. (2004). Funktionale Kompetenz oder mathematische Allgemeinbildung?: Thesen zu "(Nationale) Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss". In J. Schlömerkemper (Hrsg.), *Bildung und Standards: Zur Kritik der "Instandardsetzung" des deutschen Bildungswesens. Die Deutsche Schule*, 8. Beiheft, S. 135-144.
- Jäger, R. S., Frey, A. & Wosnitza, M. (Hrsg.). *Lernprozesse, Lernumgebungen und Lern Diagnostik. Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert*. Landau: Verlag Empirische Pädagogik.
- Jansen, M., Schroeders, U., & Stanat, P. (2013). Motivationale Schülermerkmale in Mathematik und den Naturwissenschaften. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathema-*

- tische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I (S. 347–366). Münster: Waxmann.
- Kaiser, G. & Henn, H.-W. (Hrsg.). (2015). *Werner Blum und seine Beiträge zum Modellieren im Mathematikunterricht. Festschrift zum 70. Geburtstag von Werner Blum*. Wiesbaden: Springer Spectrum.
- Kenward, M. G. & Carpenter, J. R. (2009). Multiple Imputation. In G. Fitzmaurice, M. Davidian, G. Verbeke & G. Molenberghs (Hrsg.), *Longitudinal Data Analysis* (S. 477-501). Boca Raton: CRC Press.
- Kiplinger, V. L., Haug, C. A. & Ademi, J. (2000). *Measuring Math – not Reading – on a Math Assessment: A Language Accomodations Study of English Language Learners and Other Special Populations*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. April 24-28, 2000.
- Klafki, Wolfgang (1985). *Neue Studien zur Bildungstheorie und Didaktik*. Weinheim: Beltz.
- Klauer, K. J. (Hrsg.). (2001) *Handbuch kognitives Training*. Göttingen: Hogrefe.
- Klein, H. P. & Jahnke, T. (2012). Die Folgen der Kompetenzorientierung im Fach Mathematik. *Journal der Didaktik der Biowissenschaften*, 2, S. 1-9.
- Klieme, E. & Leutner D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52 (6), S. 876-903.
- Klieme, E. & Reusser, K. (2003). Unterrichtsqualität und mathematisches Verständnis im internationalen Vergleich? Ein Forschungsprojekt und erste Schritte zur Realisierung. *Unterrichtswissenschaft*, 31(3), 194-205.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M, Reiss, K., Ri-quarts, K., Rost, J, Tenorth, H.-E., Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“.In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127-146). Münster: Waxmann.

- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43-58). Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Klingberg, L. (1986). Unterrichtsprozess und didaktische Fragestellung. Berlin.
- Klinger, Udo (Hrsg.), *Mit Kompetenz Unterricht entwickeln; Fortbildungskonzepte und –materialien*. Troisdorf: Bildungsv Verlag Eins.
- KMK (2003a). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Beschluss der Kultusministerkonferenz vom 04.12.2003*. Luchterhand.
- KMK (2003b): *Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10)*
[Online:
http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mittleren-SA.pdf. 14.02.2014]
- KMK (2006): *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*
[Online: :
http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_08_01-Gesamtstrategie-Bildungsmonitoring.pdf. 14.02.2014]
- KMK (2012). *Kompetenzstufenmodell zu den Bildungsstandards für den Hauptschulabschluss und den Mittleren Schulabschluss im Fach Mathematik*.
<http://www.iqb.hu-berlin.de/bista/ksm> [Zugriff am 22.03.2014].
- Knapp, M. S. (1997). Between Systemic Reforms and the Mathematics and Science Classroom: The Dynamics of Innovation, Implementation, and Professional Learning. *Review of Educational Research*, 67 (2), S. 227-266.
- Knapp, M. S. (2002). Understanding how policy meets practice: Two takes on local responses to a state reform initiative. Seattle: University of Washington, Center for the Study of Teaching and Policy.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Köller, O. (2008). Bildungsstandards – Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik*, 54 (2), S. 163-173.
- Köller, O. (2010). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 529-548). Wiesbaden: Verlag für Sozialwissenschaften.

- Köller, O. & Baumert, J. (2012). Schulische Leistungen und ihre Messung. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie* (7. Aufl.) (S. 645–661). Weinheim: Beltz.
- Köller, O. & Karim, A. (2001). *Der Einfluss der Schule auf die Intelligenzentwicklung im Jugendalter*. Vortrag auf der 15. Tagung der Fachgruppe Entwicklungspsychologie, Potsdam.
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, 20 (1/2), S. 27-39.
- Komorek, E., Bruder, R., Collet, C. & Schmitz, B. (2006). Inhalte und Ergebnisse einer Intervention im Mathematikunterricht der Sekundarstufe I mit einem Unterrichtskonzept zur Förderung mathematischen Problemlösens und von Selbstregulationskompetenzen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 240-267). Münster: Waxmann.
- Komorek, E., Bruder, R., Schmitz, B. (2004). Integration evaluierter Trainingskonzepte für Problemlösen und Selbstregulation in den Mathematikunterricht. In J. Doll & M. Prenzel (Hrsg.), *Schulische und außerschulische Ansätze zur Verbesserung der Bildungsqualität* (S. 54-76). Münster: Waxmann.
- Kounin, J. S. (1976). *Techniken der Klassenführung*. Bern: Huber.
- Krammer, K., Schnetzler, C. L., Pauli, C., Reusser, K., Ratzka, N., Lipowsky, F. & Kliehme, E. (2010). Unterrichtsvideos in der Lehrerfortbildung. Überblick über Konzeptionen und Ergebnisse einer einjährigen netzgestützten Fortbildungsveranstaltung. In F. H. Müller, A. Eichenberger, M. Lüders & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung* (S. 227-244). Münster: Waxmann.
- Kruse, S., Louis, K.S., Bryk, A. (1995). An Emerging Framework for Analyzing School-Based Professional Community. In K. S. Louis, S. Kruse et al. (Hrsg.), *Professionalism and Community: Perspectives on Reforming Urban Schools*. Thousand Oaks (S. 23-44). CA: Corwin Press.
- Kubina, C. (Hrsg.) (2009). *Kompetenzorientiert unterrichten, Bildungsstandards nutzen*. Unterstützungsprogramm zur Unterrichtsentwicklung 2209/2010. 1. Aufl., Frankfurt am Main: Amt für Lehrerbildung.

- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. Münster: Waxmann.
- Kunter, M. & Pohlmann, B. (2009). Lehrer. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 261-282). Heidelberg: Springer Medizin Verlag.
- Kunter, M., Brunner, M., Baumert, J., Klusmann, U., Krauss, S., Blum, W., Jordan, A. & Neubrand, M. (2005). Der Mathematikunterricht der PISA-Schülerinnen und Schüler. *Zeitschrift für Erziehungswissenschaft*, 8 (4), S. 502-520.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Landesinstitut für Schule und Medien Berlin-Brandenburg (2009). *Implementation der Bildungsstandards im Fach Mathematik. Projektbericht*. Unveröffentlichter Projektbericht. Ludwigsfelde: LISUM.
- Leiß, D. & Blum, W. (2006). Beschreibung zentraler mathematischer Kompetenzen. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 33-50). Berlin: Cornelsen Scriptor.
- Lemmermöhle, D., Rothgangel, M., Bögeholz, S., Hasselhorn, M. & Watermann, R. (Hrsg.). (2007). *Professionell lehren, erfolgreich lernen*. Münster: Waxmann.
- Lenski, A. E., Richter, D. & Pant, H. A. (2015). Kompetenzorientierung im Unterricht aus der Perspektive von Lehrkräften und Schülerinnen und Schüler. *Zeitschrift für Pädagogik*, 61 (5), S. 712-737.
- Lersch, R. (2010). Didaktik und Praxis kompetenzfördernden Unterrichts. In Faulstich-Christ, K., Lersch, R. & Moegling, K. (Hrsg.), *Kompetenzorientierung in Theorie, Forschung und Praxis* S. 31—60. Immenhausen bei Kassel.
- Leuders, T. & Leiß, D. (2006). Realitätsbezüge. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 194-206). Berlin: Cornelsen Scriptor.
- Leuders, T. (2001). *Qualität im Mathematikunterricht der Sekundarstufe I und II*. Berlin: Cornelsen Scriptor.
- Leuders, T. (2006). Kompetenzorientierte Aufgaben im Unterricht. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret*.

- Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 81–95). Berlin: Cornelsen Scriptor.
- Leutner, D., Klieme E., Meyer, K. & Wirth, J. (2004). Problemlösen. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland, Ergebnisse des zweiten internationalen Vergleichs* (S. 147-175). Münster: Waxmann.
- Lipowsky, F. (2010). Lernen im Beruf. Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In F. H. Müller, A. Eichenberger, M. Lüders & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung* (S.51-72). Münster: Waxmann.
- Lipowsy, F. (2004). Was macht Fortbildungen für Lehrkräfte erfolgreich? *Die Deutsche Schule*, 96, H. 4, S. 462-479.
- Little, J. W. (2007). Teachers accounts of classroom experience as a resource for professional learning and instructional decision making. *Yearbook of the National Society for the Study of Education*, 106, S. 217-240.
- Little, J. W.& McLaughlin, M.W. (1993). *Teachers' work: Individuals, colleagues, and contexts*. New York: Teachers College Press.
- Maag Merki, K. (2004). Lernkompetenzen als Bildungsstandards – eine Diskussion der Umsetzungsmöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, 7 (4), S. 537-550.
- Maag Merki, K. (Hrsg.) (2009). *Kooperation und Netzwerkbildung. Strategien zur Qualitätsentwicklung in Schulen*. Seelze: Klett Kallmeyer.
- Maag Merki, K. (2010). Theoretische und empirische Analysen zur Effektivität von Bildungsstandards, standardbezogenen Lernstandserhebungen und zentralen Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung im Schulsystem* (S. 145-170). Wiesbaden: Verlag für Sozialwissenschaften.
- Maag Merki, K. (2016). Theoretische und empirische Analysen der Effektivität von Bildungsstandards, standardbezogene Lernstandserhebungen und zentralen Abschlussprüfungen. In H. Altrichter & K. Maag Merki (Hrsg.), *Handbuch Neue Steuerung in Schulsystem 2. Auflage* (S. 151-181). Wiesbaden: Springer VS.

- Maag Merki, K. & Schwippert, K. (2008). Systeme der Rechenschaftslegung und Schulentwicklung: Editorial. *Zeitschrift für Pädagogik*, 54 (6), S. 773-776.
- Martinek, D. (2012). Autonomie und Druck im Lehrberuf. *Zeitschrift für Bildungsforschung*, 2. Jg., S. 23-40.
- Mason, B., Mason, D.W., Mendez, M., Nelsen, G. & Orwig, R. (2005). Effects of Top-Down and Bottom-Up Elementary School Standards Reform in an Underperforming California District. *Elementary School Journal* , 105 (4), S. 353-376.
- Massell, D., & Consortium for Policy Research in Education. (1998). *State strategies for building capacity in education: Progress and continuing challenges* (CPRE Research Report Series RR-41). Philadelphia: Author.
- Mayer, R. E. (2004). Should There Be a Three Strikes-Rule Against Pure Discovery Learning? The Case for Guided Methods of Instruction. *American Psychologist*, 59 (1), S. 14-19.
- Menard, S. (2002). *Applied Logistic Regression Analysis. 2nd edition*. Sage University Papers Series on Quantitative Applications in the Social Sciences, 106. Thousand Oaks, CA: Sage.
- Messner, R. (2009). Bausteine eines kognitiv-aktivierenden Fachunterrichts. In D. Bosse (Hrsg.), *Gymnasiale Bildung zwischen Kompetenzorientierung und Kulturarbeit* (S. 137-160). Wiesbaden: Verlag für Sozialwissenschaften.
- Mitchell, C. & Sackney, L. (2000). *Profound Improvement. Building Capacity for a Learning Community*. Lisse: Swets&Zeitlinger.
- Moegling, K. (2010). Die Kompetenzdebatte – Zum Verhältnis von Bildung und Kompetenzorientierung. In Faulstich-Christ, K., Lersch, R. & Moegling, K. (Hrsg.), *Kompetenzorientierung in Theorie, Forschung und Praxis* S. 4-30. Immenhausen bei Kassel.
- Müller, F. H., Eichenberger, A., Lüders, M. & Mayr, J. (Hrsg.). (2010). *Lehrerinnen und Lehrer lernen. Konzepte und Befunde zur Lehrerfortbildung*. Münster: Waxmann.
- National Commission on Excellence in Education (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: U.S. Government Printing Office.
- National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- Neubrand, M. (2006). Multiple Lösungswege für Aufgaben: Bedeutung für Fach, Lernen, Unterricht und Leistungserfassung. In W. Blum, C. Drüke-Noe, R. Hartung & O.

- Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 162-177). Berlin: Cornelsen Scriptor.
- Newman, F.M. (1994). School-wide Professional Community. *Issues in Restructuring Schools*, No. 6, S. 1-3.
- Newman, F.M. et al. (1996). *Authentic Achievement – Restructuring Schools for Intellectual Quality*. San Francisco: Jossey-Bass Publishers.
- Newman, F., King, B., & Youngs, P. (2000). Professional development that addresses school capacity: Lessons from urban elementary schools. *American Journal of Education*, 108(4), 259-299.
- Oelkers, J. & Reusser, K. (2008). *Expertise: Qualität entwickeln – Standards sichern – mit Differenz umgehen*. Bonn, Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Ostermeier, C., Carstensen, C. H., Prenzel, M. & Geiser, H. (2004). Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken. Ausgangsbedingungen für die Implementation im BLK-Modellversuchsprogramm SINUS. *Unterrichtswissenschaft*, 32. Jg., H. 3, S. 215-237.
- Otto, B., Perels, F., Schmitz, B. & Bruder, R. (2006). Längsschnittliche und prozessuale Evaluation eines Trainingsprogramms zur Förderung sachspezifischer und fächerübergreifender (selbstregulativer) Kompetenzen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 211-239). Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008a). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 54. Jg., Heft 6, S. 827-845.
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008b). Eine modellbasierte Erfassung der Auseinandersetzung von Lehrkräften mit den länderübergreifenden Bildungsstandards. In E.-M. Lankes (Hrsg.), *Pädagogische Professionalität als Gegenstand Empirischer Forschung* (S. 245-260). Münster: Waxmann.

- Pauli, C, Drollinger-Vetter, B, Hugener, I. & Lipowsky, F. (2008). Kognitive Aktivierung im Mathematikunterricht. *Zeitschrift für Pädagogische Psychologie*, 22 (2), S. 127-133.
- Pauli, C. & Reusser, K. (2010). Selbst- und Unterrichtswahrnehmung von Lehrpersonen. In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 143-170). Münster: Waxmann.
- Peek, R. (2004). Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM). Klassenbezogene Ergebnisrückmeldungen und ihre Rezeption in Brandenburger Schulen. In B. Kohler, F. W. Schrader (Hrsg.), *Ergebnisrückmeldung und Rezeption. Empirische Pädagogik 18 [Themenheft]* (S. 82-114). Landau: Verlag Empirische Pädagogik.
- Pekrun, R., vom Hofe, R., Blum, W., Götz, T., Wartha, S., Frenzel, A. & Jullien, S. (2006). Projekt zur Analyse der Leistungsentwicklung in Mathematik (PALMA). Entwicklungsverläufe, Schülervoraussetzungen und Kontextbedingungen von Mathematikleistungen in der Sekundarstufe I. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule* (S. 21-53). Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann.
- Pekrun, R. & Zirngibl, A. (2004). Schülermerkmale im Fach Mathematik. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-J. Rolff, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 191-210). Münster: Waxmann.
- Pelletier, L.G., Séguin-Lévesque, C. & Legault, L. (2002). Pressure from above and pressure from below as determinants of teacher's motivation and teaching behaviors. *Journal of Educational Psychology*, 94 (1), S. 186-196.
- Pöhlmann, C., Pant, H. A., Frenzel, J., Roppelt, A. & Köller, O. (2014). Auswirkungen einer Intervention auf die Auseinandersetzung und Arbeit mit Bildungsstandards bei Mathematik-Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 17 (1), 2014, S. 113-133.
- Popham, W. J. (2004). *America's „Failing“ Schools*. New York: Routledge Falmer.
- Porter, A. C., Floden, R., Freeman, D., Schmidt, W., & Schwille, J. (1988). Content determinants in elementary school mathematics. In D. Gouws & T. Cooney (Hrsg.),

- Perspectives on research on effective mathematics teaching*. Reston: National Council of Teachers of mathematics.
- Prenzel, M. (2005). Bildungsstandards und die Weiterentwicklung des Mathematikunterrichts. In: H.-W. Henn, & G. Kaiser (Hrsg.), *Mathematikunterricht im Spannungsfeld von Evolution und Evaluation. Festschrift für Werner Blum* (S. 261–271). Hildesheim: Franzbecker.
- Prenzel, M. & Allolio-Näcke, L. (Hrsg.) (2006). Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms. Münster: Waxmann.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. & Schiefele, U. (Hrsg.). (2005). *PISA 2003. Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche?* Münster: Waxmann.
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost J. & Schiefele, U. (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland, Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Prenzel, M., Brackhahn, B. & Hertrampf, M. (2002). *Konzeption zur Dissemination des BLK-Modellversuchsprogramms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“*. Kiel: IPN.
- Prenzel, M., Walter, O & Frey, A. (2007). PISA misst Kompetenzen. Eine Replik auf Rindermann (2006): Was messen internationale Schulleistungstudien. *Psychologische Rundschau*, 58 (2), S. 128-136.
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J. Schiefele, U. (2006). *PISA 2003 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd edition*. Newbury Park, CA: Sage.
- Ravitch, D. (1995). *National standards in American education. A citizen's guide*. Washington DC: Brookings Institution.
- Reeves, D. (2006). *The learning leader*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Reinisch, H. (2009). "Lehrerprofessionalität" als theoretischer Term - Eine begriffssystematische Analyse. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität - Bedingungen, Genese, Wirkungen und Messung* (S. 33-44). Weinheim: Beltz.
- Reinmann-Rothmeier, G. & Mandl, H. (2001). Wissensmanagement. *Psychologie in Erziehung und Unterricht*, 1, S. 76-77.
- Reiss, K. (2004). Bildungsstandards und die Rolle der Fachdidaktik am Beispiel der Mathematik. *Zeitschrift für Pädagogik*, 50 (5), S. 635-649.
- Reiss, K. & Reiss, M. (2006). Unterrichtsqualität und der Mathematikunterricht. In I. Hosenfeld & F.W. Schrader (Hrsg.), *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven*. (S. 225-242). Münster: Waxmann.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos (Hrsg.), *Bildungsstandards Deutsch und Mathematik - Leistungsmessung in der Grundschule* (S. 120-141). Weinheim/Basel: Beltz.
- Reiss, K., Heinze, A., Kuntze, S., Kessler, S., Rudolph-Albert, F. & Renkl, A. (2006). Mathematiklernen mit heuristischen Lösungsbeispielen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 194-210). Münster: Waxmann.
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs, P. Füglistner, K. Reusser & H. Wyss (Hrsg.), *Didaktik auf psychologischer Grundlage. Von Hans Aebli's kognitionspsychologischer Didaktik zur modernen Lehr-Lernforschung* (S. 151-168). Bern: h.e.p.
- Reusser, K., Pauli, C. & Elmer, A. (2011). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 478-496). Münster: Waxmann.
- Reusser, K., Pauli, C. & Waldis, M. (Hrsg.) (2010). *Unterrichtsgestaltung und Unterrichtsqualität. Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht*. Münster: Waxmann.
- Rindermann, H. (2006). Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz? *Psychologische Rundschau*, 57 (2), S. 69-86.

- Rindermann, H. (2007). Intelligenz, kognitive Fähigkeiten, Humankapital und Rationalität auf verschiedene Ebenen. [Antwort auf die Kritiken von Prenzel et al. und Baumert et al. in der *Psychologischen Rundschau* auf den Beitrag „Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz?“ in der *Psychologischen Rundschau* 2006, 57(2)]. *Psychologische Rundschau*, 58 (2), S. 137-145.
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). New York: Free Press.
- Rokitta, J., Meisner, A. & Hirschmann, I. (2009). Die Bildungsstandards in der schulischen Realität. In R. Wernstedt & M. John-Ohnesorg (Hrsg.), *Bildungsstandards als Instrument schulischer Qualitätsentwicklung. Zementierung des Selektionsprinzips oder Mittel zur Chancengerechtigkeit?* (S. 34-38). Berlin: bub.
- Rolff, H.-G. (2001): Schulentwicklung konkret. Steuergruppe, Bestandsaufnahme, Evaluation. Velber: Kallmeyersche Verlagsbuchhandlung.
- Roppelt, A., Penk, C., Pöhlmann, C. & Pietsch, E. (2013). Der Ländervergleich im Fach Mathematik. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. (S. 123-140). Münster: Waxmann.
- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Rost, J. (2004). Psychometrische Modelle zur Überprüfung von Bildungsstandards anhand von Kompetenzmodellen. *Zeitschrift für Pädagogik*, 50 (5), S. 662-678.
- Rost, D. H. (2005). *Interpretation und Bewertung pädagogisch-psychologischer Studien*. Weinheim: Beltz.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63 (3), S. 581-592.
- Rudow, B. (1994). *Die Arbeit des Lehrers. Zur Psychologie der Lehrertätigkeit, Lehrbelastung und Lehrergesundheit*. Bern: Hans Huber.
- Rupp, A. A., Leucht, M. & Hartung, R. (2006). „Die Kompetenzbrille aufsetzen“. Verfahren zur multiplen Klassifikation von Lernenden für Kompetenzdiagnostik in Unterricht und Testung. *Unterrichtswissenschaft*, 34 (3), S. 195-219.
- Schecker, H. (2012). Standards, Competencies and Outcomes. A Critical View. In S. Bernholt, K. Neumann & P. Nentwig (Hrsg.), *Making It Tangible - Learning Outcomes in Science Education* (S. 237-253). Münster: Waxmann.

- Schellenbach-Zell, J., Rürup, M., Fussangel, K. & Gräsel, C. (2008). Bedingungen erfolgreichen Transfers am Beispiel von Chemie im Kontext. In R. Demuth, C. Gräsel, B. Ralle & I. Parchmann (Hrsg.), *Chemie im Kontext - Von der Innovation zur nachhaltigen Verbreitung einer Unterrichtskonzeption* (S. 83-123). Münster: Waxmann.
- Schott, F. & Ghanbari, A. (2008). *Kompetenzdiagnostik, Kompetenzmodelle, kompetenzorientierter Unterricht. Zur Theorie und Praxis überprüfbarer Bildungsstandards*. Münster: Waxmann.
- Schreder, G. (2007): Wirksames Lernen in der Lehrerfortbildung. Lehrerinnen und Lehrer verändern ihren Unterricht. *Die Deutsche Schule*, 99, H. 4, S. 480-488.
- Singer, J. D. & Willet, J. B. (2003). *Applied Longitudinal Data Analysis. Modelling Change and Event Occurance*. New York: Oxford University Press.
- Sjuts, J. (2006). Unterrichtliche Gestaltung und Nutzung kompetenzorientierter Aufgaben in diagnostischer Hinsicht. In: W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 96-112) . Berlin: Cornelsen Scriptor.
- Spillane, J. (2004). *Standards Deviation: How schools misunderstand education policy*. Cambridge, MA., Harvard University Press.
- Spillane, J. P., Reiser, B. J. & Reimer, T. (2002). Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. *Review of Educational Research*, 72 (3), S. 387-431.
- Stadler, M., Ostermeier, C. & Prenzel, M. (2005). *Bericht zur ersten Verbreitungswelle 2003-2005*. BLK-Programm SINUS-Transfer. [Zugriff am 18.03.2017 auf http://www.sinus-transfer.de/fileadmin/MaterialienIPN/Bericht_1W_IPN.pdf]
- Staub F. C. & Stern, E. (2002): The Nature of Teachers' Pedagogical Content eliefs Matters for Students' Achievement Gains: Quasi-Experimental Evidence From Elementary Mathematics. In: *Journal of Educational Psychology*, 94/2, S. 344-355.
- Stipek, D. J., Givvin, K. B., Salmon, J. M. & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 17. Jg, S. 213-226.
- Sjuts, J. (2006). Unterrichtliche Gestaltung und Nutzung kompetenzorientierter Aufgaben in diagnostischer Hinsicht. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller

- (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 96-112). Berlin: Cornelsen Scriptor.
- Stoll, L. (2009). Capacity building for school improvement or creating capacity for learning? A changing landscape. In *Journal of Educational Change*, 10 (2), S. 115-127.
- Stoll, L., Bolam, R., McMahon, A., Wallace, M. & Thomas, S. (2006). Professional Learning Communities: a Review of Literature. *Journal of Educational Change*, 7, S. 221-258.
- Sullivan, P. & Wood, T. (Hrsg.) (2008), *International handbook of mathematics teacher education*. Vol.1: Knowledge and beliefs in mathematics teaching and teaching development. Rotterdam: Sense Publishers.
- Terhart, E., Bennewitz, H. & Rothland, M. (Hrsg.). (2011). *Handbuch der Forschung zum Lehrerberuf*. Münster: Waxmann.
- Tiffin-Richards, S. (2011). Setting Standards for the Assessment of English as a Foreign Language: Establishing validity evidence for criterion-referenced interpretations of test-scores. Dissertation Freie Universität Berlin.
- Tippelt, R. & Schmidt, B. (Hrsg.). (2010). *Handbuch Bildungsforschung*. Wiesbaden: Verlag für Sozialwissenschaften.
- Trautwein, U., Köller, O. & Baumert, J. (2001). Lieber oft als viel: Hausaufgaben und die Entwicklung von Leistung und Interesse im Mathematik-Unterricht der 7. Jahrgangsstufe. *Zeitschrift für Pädagogik*, 47 (5), S. 703-724.
- Van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research*, 72, S. 577-625.
- Wackermann, R. (2008). *Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer*. Berlin: Logos.
- Walter, O. (2005). *Kompetenzmessung in den PISA-Studien. Simulation zur Schätzung von Verteilungsparametern und Reliabilitäten*. Lengerich: Pabst.
- Weinert, F. E. (1998). Neue Unterrichtskonzepte zwischen gesellschaftlichen Notwendigkeiten, pädagogischen Visionen und psychologischen Möglichkeiten. In Bayerisches Staatsministerium für Unterricht, Kultur, Wissenschaft und Kunst (Hrsg.), *Wissen und Werte für die Welt von morgen* (S. 104-125). München.

- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In Weinert, F. E. (Hrsg.), *Leistungsmessung in Schulen* (S. 17-31). Weinheim: Beltz.
- Weinert, F. E. (Hrsg.). (2001). *Leistungsmessung in Schulen*. Weinheim: Beltz Verlag, S. 17-31.
- Weinert, F.E., Schrader, F.W. & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13, S. S. 895–914.
- Winkelmann, H. & Robitzsch, A. (2009). Modelle mathematischer Kompetenzen: Empirische Befunde zur Dimensionalität. In: D. Granzer, O. Köller, A. Bremerich-Vos (Hrsg.), *Bildungsstandards Deutsch und Mathematik - Leistungsmessung in der Grundschule* (S. 169-197). Weinheim/Basel: Beltz.
- Winkelmann, H., Robitzsch, A., Stanat, P. & Köller, O. (2012). Mathematische Kompetenzen in der Grundschule. Struktur, Validierung und Zusammenspiel mit allgemeinen kognitiven Fähigkeiten. *Diagnostica*, 58 (1), S. 15-30.
- Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, S. 37-46.
- Winter, H. (2003). Mathematikunterricht und Allgemeinbildung. In H.-W. Henn & K. Maaß (Hrsg.), *Materialien für einen realitätsbezogenen Mathematikunterricht* (S. 6-15). Hildesheim: Franzbecker. Ursprünglich publiziert in: *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61/1995, 37-46.
- Wu, M. L., Adams, R. J. & Wilson, M. R. (1998). ConQuest: Generalized item response modelling software. Melbourne: Australian Council for Educational Research.
- Wynands, A. (2006). Intelligentes Üben. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen* (S. 113-125). Berlin: Cornelsen Scriptor.
- Zeitler, S., Asbrand, B. & Pöhlmann, C. (2009). Unterrichtsentwicklung durch unterstützte Kooperation in Fachgruppen? Ein Projekt zur Implementation der Bildungsstandards in Berliner und Brandenburger Schulen. In K. Maag Merki (Hrsg.), *Kooperation und Netzwerkbildung. Strategien zur Qualitätsentwicklung in Schulen* (S. 14-25). Seelze: Klett Kallmeyer.

- Zeitler, S., Asbrand, B. & Tesch, B. (2010). Bildungsstandards und ihre Implikationen für Qualitätssicherung und Qualitätsentwicklung. In A. Gehrmann, U. Hericks & M. Lüders (Hrsg.), *Bildungsstandards und Kompetenzmodelle. Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (S. 15-28). Bad Heilbrunn: Klinkhardt.
- Zeitler, S., Heller, N. & Asbrand, B. (2012). *Bildungsstandards in der Schule. Eine rekonstruktive Studie zur Implementation der Bildungsstandards*. Münster: Waxmann.
- Ziener, G. (2006). *Bildungsstandards in der Praxis – Kompetenzorientiert unterrichten*. Seelze-Velber: Kallmeyer.
- Zimmer, K. Brunner, M., Lüdtke, O., Prenzel, M. & Baumert J. (2007). Die PISA-Spitzengruppe in Deutschland: Eine Charakterisierung hochkompetenter Jugendlicher. In K. A. Heller & A. Ziegler (Hrsg.), *Begabt sein in Deutschland* (S. 193-208). Münster: Lit.
- Zlatkin-Troitschanskaia, O. & Seidel, J. (2011). Kompetenz und ihre Erfassung – das neue „Theorie-Empirie-Problem“ der empirischen Bildungsforschung?. In O. Zlatkin-Troitschanskaia (Hrsg.), *Stationen empirischer Bildungsforschung*, S. 218-233. Wiesbaden: Verlag für Sozialwissenschaften.

9 Abbildungsverzeichnis

Abbildung 1.	Kompetenz und Performanz aus didaktischer Perspektive nach Lersch, 2010.....	27
Abbildung 2.	Kompetenzerwerb nach Lersch (2010).	29
Abbildung 3.	Kompetenzmodell für das Fach Mathematik, Sekundarstufe I, Quelle: Blum et al. (2006).	38
Abbildung 4.	Die Aufgabe „Näherungswert für die Kreiszahl π “ (Quelle: Blum et al., 2006, S. 218).	41
Abbildung 5.	Lage von Mindest- und Regelstandards der HSA- und MSA-Population auf dem integrierten Kompetenzstufenmodell Mathematik, Quelle: KMK, 2012.	44
Abbildung 6.	Gegenüberstellung von altem und neuem Unterrichtsskript, Quelle: Lersch, 2010.	55
Abbildung 7.	Aufgabe "Regel", Quelle: IQB.	75
Abbildung 8.	Aufgabe "Raute", Quelle: Neubrand (2006, S. 174).	76
Abbildung 9.	Das Concern-Based-Adoption-Model (Hall & Hord, 2006, S. 108).	94
Abbildung 10.	Verbreitung der SINUS-Konzeption über die Zeit, Quelle: http://mathematik.bildung-rp.de/sekundarstufe-i/thema/sinus-und-sinus-transfer/sinus-transfer-3-2007-2009.html [Zugriff am 23.11.2013].	125
Abbildung 11.	Untersuchungsdesign der quantitativ-standardisierten Begleitforschung im Zeitraum 2007-2009; Quelle: eigene Darstellung.	135
Abbildung 12.	Projektdauer und teilnehmende Schulen im Projekt "Erprobung eines Konzepts zur Implementation der Bildungsstandards im Fach Mathematik in Berliner und Brandenburger Schulen"; Quelle: eigene Darstellung.	143
Abbildung 13.	Stichprobe der Lehrkräfte der Studie 1.	146
Abbildung 14.	Übersicht über die Module im Projekt "Kompetenzorientiert unterrichten in Mathematik und Naturwissenschaften"; Quelle: Kubina, 2009.	151
Abbildung 15.	Design des Mathematiktests.	165
Abbildung 16.	Positive Einstellungen gegenüber den Bildungsstandards im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	185
Abbildung 17.	Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der positiven Einstellungen gegenüber den Bildungsstandards.	186
Abbildung 18.	Negative Einstellungen gegenüber den Bildungsstandards im Verlauf; Gegenüberstellung gesamte Stichprobe und Längsschnitt.	188
Abbildung 19.	Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der negativen Einstellungen gegenüber den Bildungsstandards.	189

Abbildung 20. Zusammenarbeit in der Fachkonferenz im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	193
Abbildung 21. Intraindividuelle Entwicklung der Lehrkräfte (Studie 1) für das Merkmal der Zusammenarbeit in der Fachkonferenz.	194
Abbildung 22. Kooperation in der Fachkonferenz im Verlauf; Gegenüberstellung Querschnitt, Längsschnitt und imputierte Daten.	196
Abbildung 23. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal Kooperation in der Fachkonferenz.	197
Abbildung 24. Mathematisches Argumentieren im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	200
Abbildung 25. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisches Argumentieren.	201
Abbildung 26. Probleme mathematisch Lösen im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	203
Abbildung 27. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal Probleme mathematisch Lösen.	204
Abbildung 28. Mathematisches Modellieren im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	206
Abbildung 29. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisches Modellieren.	207
Abbildung 30. Mathematische Darstellungen Verwenden im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	208
Abbildung 31. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematische Darstellungen Verwenden.	209
Abbildung 32. Technisch Arbeiten im Verlauf; Gegenüberstellung gesamte Stichprobe, Längsschnitt und imputierter Datensatz.	210
Abbildung 33. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal technisch Arbeiten.	211
Abbildung 34. Mathematisches Kommunizieren im Verlauf, Gegenüberstellung von Querschnitt, Längsschnitt und imputierten Daten.	213
Abbildung 35. Intraindividuelle Entwicklung der Lehrkräfte im Merkmal mathematisch kommunizieren.	213
Abbildung 36. Positive Einstellungen gegenüber den Bildungsstandards - Veränderungen; Gegenüberstellung global und differenziert.	218
Abbildung 37. Negative Einstellungen gegenüber den Bildungsstandards - Veränderungen; Gegenüberstellung global und differenziert.	220
Abbildung 38. Kooperation in der Fachkonferenz - Veränderungen; Gegenüberstellung global und differenziert.	222

Abbildung 39. Zusammenarbeit in der Fachkonferenz - Veränderung; Gegenüberstellung global und differenziert.	223
Abbildung 40. Mathematisch Argumentieren - Veränderung; Gegenüberstellung global und differenziert.	225
Abbildung 41. Probleme mathematisch Lösen - Veränderung; Gegenüberstellung global und differenziert.	227
Abbildung 42. Mathematisch Modellieren - Veränderung; Gegenüberstellung global und differenziert.	228
Abbildung 43. Mathematische Darstellungen Verwenden - Veränderung; Gegenüberstellung global und differenziert.	230
Abbildung 44. Technisch Arbeiten - Veränderung; Gegenüberstellung global und differenziert.	231
Abbildung 45. Mathematisch Kommunizieren - Veränderung; Gegenüberstellung global und differenziert.	232
Abbildung 46. Mathematikleistung im Verlauf - Gegenüberstellung aller Teilnehmer, Längsschnitt und imputierter Datensatz.	260
Abbildung 47. Veränderung in der Mathematikleistung, Globaldarstellung und detaillierte Darstellung.	272
Abbildung 48. Antworthäufigkeiten in Prozent zum Item: "Die professionelle Handlungs- und Entscheidungsfreiheit wird durch Standards deutlich eingeschränkt."	281
Abbildung 49. Antworthäufigkeiten in Prozent zum Item: "Am ganzen Standard-Konzept ist noch zu Vieles ungeklärt, um es wirklich beurteilen zu können."	282

10 Tabellenverzeichnis

Tabelle 1:	Verteilung der Schülerinnen und Schüler der 9. und 10. Jahrgangsstufe auf die Kompetenzstufen, Quelle: KMK, 2012, eigene Darstellung.	48
Tabelle 2:	Merkmale von neuem und altem Unterrichtsskript, eigene Darstellung	59
Tabelle 3:	Zentrale Veranstaltungen im Projekt zur "Erprobung eines Fortbildungskonzeptes in Berliner und Brandenburger Schulen"	137
Tabelle 4:	Übersicht der Schulen und Klassen der Studie 1	145
Tabelle 5:	Schulentwicklungserfahrung und Zuordnung im Experimentaldesign der hessischen Schulen.....	153
Tabelle 6:	Übersicht der Schulen und Klassen in Studie 2	154
Tabelle 7:	Beispielitems für die Einstellungsskalen zur Einführung der nationalen Bildungsstandards	156
Tabelle 8:	Reliabilitätsmaße der Skalen zu den Einstellungen gegenüber den Bildungsstandards (Lehrkräfte).....	157
Tabelle 9:	Beispielitems zu den Skalen Zusammenarbeit in der Fachkonferenz und Kooperation	157
Tabelle 10:	Reliabilitätsmaße der Skalen „Zusammenarbeit in der Fachkonferenz“ und „Kooperation in der Fachkonferenz“ (Lehrkräfte).....	158
Tabelle 11:	Beispielitems zum Fragebogen "Kompetenzorientierte Schülertätigkeiten"	159
Tabelle 12:	Reliabilitätsmaße zu den Skalen der kompetenzorientierten Tätigkeiten (Lehrkräfte)	160
Tabelle 13:	Reliabilitätsmaße zu den Skalen der kompetenzorientierten Tätigkeiten (Schülerinnen und Schüler)	160
Tabelle 14:	Beispielitems zu den Skalen "mathematisches Selbstkonzept" und "Interesse an Mathematik" im Schülerfragebogen.....	162
Tabelle 15:	Reliabilitätsmaße zu den Skalen „mathematisches Selbstkonzept“ und „Interesse an Mathematik“ (Schülerinnen und Schüler)	162
Tabelle 16:	Beispielitem zur Skala "Lehrerwahrnehmung"	163
Tabelle 17:	Reliabilitätsmaße der Skala „Lehrerwahrnehmung“ (Schülerinnen und Schüler) ..	163
Tabelle 18:	Gütekriterien zum Mathematiktest der Schülerinnen und Schüler.....	166
Tabelle 19:	Missinganteile in Prozent in der Studie 1 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Lehrkräfte	171
Tabelle 20:	Missinganteile in Prozent in der Studie 1 nach Untersuchungsmerkmal und Erhebungszeitpunkt, Schüler.....	172

Tabelle 21. Missinganteile in Prozent in der Studie 2 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Lehrkräfte.....	173
Tabelle 22. Missinganteile in Prozent in der Studie 2 nach Untersuchungsmerkmal und Erhebungszeitpunkt; Schüler	174
Tabelle 23. Deskriptive Statistiken (in Prozent) für die Lehrergruppen Teilnehmer und Abbrecher in Studie 1 , Lehrkräfte.....	177
Tabelle 24. Mittelwertunterschiede zwischen den Abbrechern und Teilnehmern, Lehrerstichprobe	178
Tabelle 25. Deskriptive Statistiken (in Prozent) für die Schülergruppe der Studie 1	179
Tabelle 26. Mittelwertunterschiede zwischen Teilnehmern und Abbrechern in Studie 1, Schülerstichprobe	180
Tabelle 27. Deskriptive Statistiken (in Prozent) für die Lehrergruppen Teilnehmer und Abbrecher in Studie 2	181
Tabelle 28. Mittelwertunterschiede zwischen Teilnehmern und Abbrechern in Studie 2, Schülerstichprobe	182
Tabelle 29. Teilnehmeranzahl der Lehrkräfte über die Messzeitpunkte	183
Tabelle 30. Deskriptive Statistiken zum Merkmal positive Einstellungen gegenüber den Bildungsstandards.....	184
Tabelle 31. Deskriptive Statistik für den individuellen Wachstumsparameter aus individuellen Regressionen für positive Einstellungen gegenüber den Bildungsstandards als eine Funktion der Zeit (n=42)	186
Tabelle 32. Deskriptive Statistiken zu dem Merkmal negative Einstellungen gegenüber den Bildungsstandards.....	187
Tabelle 33. Deskriptive Statistik für den individuellen Wachstumsparameter aus individuellen Regressionen für negative Einstellungen gegenüber den Bildungsstandards als eine Funktion der Zeit (n=43)	189
Tabelle 34. Wachstumseffekte (Modell C) der Einstellungen gegenüber der Einführung von nationalen Bildungsstandards, Studie 1, Lehrkräfte	190
Tabelle 35. Deskriptive Statistiken zum Merkmal Zusammenarbeit in der Fachkonferenz.....	191
Tabelle 36. Deskriptive Statistik zum Merkmal Kooperation in der Fachkonferenz	195
Tabelle 37. Wachstumseffekte (Modell C) der Zusammenarbeit und Kooperation in der Fachkonferenz, Studie 1, Lehrkräfte	198
Tabelle 38. Deskriptive Statistiken zum Merkmal des mathematischen Argumentierens.....	199
Tabelle 39. Deskriptive Statistiken zum Merkmal des mathematischen Probleme lösens	202
Tabelle 40. Deskriptive Statistiken zum Merkmal mathematisches Modellieren	205
Tabelle 41. Deskriptive Statistiken zum Merkmal mathematische Darstellungen verwenden ...	207
Tabelle 42. Deskriptive Statistiken zum Merkmal technisch arbeiten	209

Tabelle 43.	Deskriptive Statistiken zum Merkmal mathematisch kommunizieren.....	212
Tabelle 44.	Wachstumseffekte (Modell C) der mathematischen Kompetenzen, Studie 1, Lehrkräfte.....	215
Tabelle 45.	Deskriptive Statistiken zum Merkmal "positive Einstellungen gegenüber den Bildungsstandards"	217
Tabelle 46.	Deskriptive Statistiken zum Merkmal "negative Einstellungen gegenüber den Bildungsstandards"	219
Tabelle 47.	Wachstumseffekte (Modell D) der Einstellungen gegenüber der Einführung von nationalen Bildungsstandards, Studie 2, Lehrkräfte	220
Tabelle 48.	Deskriptive Statistiken zum Merkmal "Kooperation in der Fachkonferenz"	221
Tabelle 49.	Deskriptive Statistiken zum Merkmal "Zusammenarbeit in der Fachkonferenz" ..	222
Tabelle 50.	Wachstumseffekte (Modell D) in der Zusammenarbeit in der Fachkonferenz und der Kooperation im Team, Studie 2, Lehrkräfte	224
Tabelle 51.	Deskriptive Statistiken zum Merkmal "mathematisch Argumentieren"	225
Tabelle 52.	Deskriptive Statistiken zum Merkmal "mathematisch Probleme lösen"	226
Tabelle 53.	Deskriptive Statistiken zum Merkmal "mathematisch Modellieren"	228
Tabelle 54.	Deskriptive Statistiken zum Merkmal "Darstellungen verwenden"	229
Tabelle 55.	Deskriptive Statistiken zum Merkmal "technisch arbeiten"	230
Tabelle 56.	Deskriptive Statistiken zum Merkmal "mathematisch Kommunizieren"	232
Tabelle 57.	Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 2, Lehrkräfte ..	234
Tabelle 58.	Deskriptive Statistiken zum Merkmal mathematisches Selbstkonzept, Schülerstichprobe.....	238
Tabelle 59.	Deskriptive Statistiken zum Merkmal mathematisches Interesse, Schülerstichprobe	240
Tabelle 60.	Deskriptive Statistiken zum Merkmal Lehrerwahrnehmung, Schülerstichprobe	242
Tabelle 61.	Wachstumseffekte (Modell D) zum mathematischen Selbstkonzept, Interesse an Mathematik und der Lehrerbewertung, Studie 1, Schülerinnen und Schüler	243
Tabelle 62.	Deskriptive Statistiken zum Merkmal mathematisches Argumentieren, Schüler ...	246
Tabelle 63.	Deskriptive Statistiken zum Merkmal mathematische Probleme lösen, Schüler	248
Tabelle 64.	Deskriptive Statistiken zum Merkmal mathematisch Modellieren, Schüler	250
Tabelle 65.	Deskriptive Statistiken zum Merkmal Darstellungen verwenden, Schüler	252
Tabelle 66.	Deskriptive Statistiken zum Merkmal technisch arbeiten, Schüler	254
Tabelle 67.	Deskriptive Statistiken zum Merkmal mathematisch kommunizieren, Schüler	256
Tabelle 68.	Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 1, Schülerinnen und Schüler	258
Tabelle 69.	Gegenüberstellung der Mittelwerte aus den teilnehmenden Schulen mit den Mittelwerten der Bundesländer aus dem IQB-Ländervergleich 2012.....	259

Tabelle 70. Deskriptive Statistiken zum Merkmal der Mathematikleistung, Schüler	261
Tabelle 71. Wachstumseffekte in der Mathematikleistung der Schülerinnen und Schüler, Studie 1	262
Tabelle 72. Deskriptive Statistiken zum Merkmal des mathematischen. Selbstkonzeptes, Schüler, Studie 2.....	264
Tabelle 73. Deskriptive Statistiken zum Merkmal des mathematischen Interesses, Schüler, Studie 2.....	265
Tabelle 74. Deskriptive Statistiken zum Merkmal der Lehrerbewertung, Schüler, Studie 2.....	265
Tabelle 75. Wachstumseffekte zum mathematischen Selbstkonzept, Interesse an Mathematik und der Lehrerbewertung der Schülerinnen und Schüler, Studie 2	266
Tabelle 76. Wachstumseffekte von kompetenzorientierten Tätigkeiten, Studie 2, Schülerinnen und Schüler.....	270
Tabelle 77. Deskriptive Statistiken zur Mathematikleistung, Schüler	271
Tabelle 78. Wachstumseffekte in der Mathematikleistung der Schülerinnen und Schüler, Studie 2	273

11 Anhang

Darstellung der eingesetzten Fragebögen

Items im Fragebogen „Einstellungen gegenüber den Bildungsstandards“ (Lehrkräfte)

Item-name	Itemformulierung	Skalen-zuordnung
att01	Bildungsstandards sind ein nützliches Instrument zur Qualitätssicherung im Schulwesen.	positiv
att02	Bildungsstandards machen die Ziele der Schule für alle Beteiligten transparenter.	positiv
att03	Bildungsstandards sind meiner Ansicht nach primär ein Instrument der Kontrolle von Schule und Lehrern.	negativ
att04	Die professionelle Handlungs- und Entscheidungsfreiheit wird durch Standards deutlich eingeschränkt.	negativ
att05	Wenn es nicht klare Standards gibt, ist ein Schulabschluss bald nichts mehr wert.	positiv
att06	Bildungsstandards bewirken eine "Nivellierung nach unten" und vermindern die Schulqualität eher, als dass sie sie verbessern.	negativ
att07	Bildungsstandards verstärken den Leistungs- und Selektionsdruck in der Schule.	negativ
att08	Bildungsstandards geben Klarheit, worauf es in der Schule letztendlich ankommt.	positiv
att09	Standards und Schulautonomie - das geht einfach nicht zusammen.	negativ
att10	Die unterschiedlichen Verhältnisse der Schulen in Stadt und Land erfordern gemeinsame Standards.	positiv
att11	Nur wenn wir gemeinsame Bildungsstandards haben, wird der Wert eines Schulabschlusses wieder eindeutig anerkannt werden.	positiv
att12	Standards können dazu führen, dass die Übergänge auf weiterführende Schulformen objektiviert werden und es mehr Chancengleichheit gibt.	positiv
att13	Der ausschließliche Blick auf die Ergebnisse vernachlässigt die wichtige Prozessdimension des Unterrichts.	negativ
att14	Bildungsstandards schärfen den Blick der Lehrkraft für Stärken und Schwächen ihrer Schüler/innen.	positiv
att15	Am ganzen Standard-Konzept ist noch zu Vieles ungeklärt, um es wirklich beurteilen zu können.	negativ

att16	Die einseitige Betonung von Fachleistungen führt zu einer Unterschätzung der erzieherischen Leistungen der Schule.	negativ
att17	Außer Mehrarbeit für die Lehrkräfte werden die Bildungsstandards vermutlich nicht viel bringen.	negativ
att18	Schulen sind zu verschieden, um ihre Leistungen an gleichen Maßstäben zu messen.	negativ
att19	Bildungsstandards erleichtern den Einsatz gezielter Maßnahmen zur Förderung schwächerer Schüler/innen.	positiv
att20	Von Standards zu Schulrankings ist der Weg nicht weit.	negativ
att21	Bildungsstandards geben Eltern und Schüler/innen Klarheit über das, was von der Schule erwartet wird.	positiv
att22	Bildungsstandards bringen die Gefahr einer inhaltlichen Verarmung des Unterrichts mit sich.	negativ
att23	Bildungsstandards fördern die Zusammenarbeit und Koordination der Lehrkräfte eines Faches.	positiv

Items des Fragebogens "Zusammenarbeit in der Fachkonferenz" (Lehrkräfte)

Itemname	Itemformulierung „In der Fachkonferenz...“
fach01	treffen wir uns, um eine jahrgangsübergreifende Unterrichtsplanung auszuarbeiten
fach02	sprechen wir die Maßstäbe ab, nach denen wir Klassen-, Kursarbeiten bzw. Schulaufgaben bewerten
fach03	bereiten wir gemeinsam ausgewählte Unterrichtseinheiten vor
fach04	beraten wir die Möglichkeiten individueller Förderung für Schülerinnen und Schüler
fach05	sprechen wir über die Umsetzung neuer Lehrpläne, Standards und sonstiger Vorgaben
fach06	besprechen wir die Auswahl von Lehrbüchern
fach07	besprechen wir die Anschaffungen für die Mathematiksammlung
fach08	organisieren wir schulinterne Fortbildungsveranstaltungen

Items des Fragebogens "Kooperation" (Lehrkräfte)

Itemname	Bei der Zusammenarbeit in der Fachkonferenz Mathematik hatte ich den Eindruck, dass...
koop01	wir alle an einem Strang ziehen
koop02	die Arbeitsteilung gut gelingt
koop03	wir mit einer klaren Zielvorstellung an die Arbeit gehen
koop04	in der Gruppe Einigkeit über die Zielvorstellung besteht
koop05	wir uns gegenseitig behindern
koop06	ich alleine wesentlich effektiver arbeite
koop07	die Arbeit in unserer Gruppe gerecht verteilt wird
koop08	auf unseren Treffen auf konkrete Ergebnisse hingearbeitet wird
koop09	neuer Schwung in die tägliche Arbeit kommt

Items des Fragebogens „Kompetenzorientierte Schülertätigkeiten“ (Lehrkräfte und Schülerinnen und Schüler)

Itemname	Itemformulierung
<i>Skala: mathematisches Argumentieren</i>	
st01	Mathematische Gesetzmäßigkeiten selbständig verallgemeinern.
st07	Vermutungen begründet äußern.
st13	Beweise und Herleitungen im Klassengespräch nachvollziehen.
st26	Den Gedankengang erklären, der hinter einer Idee steht.
st32	Verschiedene Argumentationen gegeneinander abwägen.
<i>Skala: Probleme lösen</i>	
st02	Selbst formulierte Fragestellungen bearbeiten.
st08	Geeignete Hilfsmittel und Strategien zum Lösen von Aufgaben finden.
st14	Mehrere Wege zur Lösung einer Aufgabe diskutieren.
st21	Aufgaben bearbeiten, bei denen nicht sofort der Lösungsweg zu erkennen ist.
st27	An Aufgaben mit "gesundem Menschenverstand" herangehen.
<i>Skala: mathematisch Modellieren</i>	
st09	Reale Situationen mit mathematischen Begriffen beschreiben.
st15	Mathematische Strukturen in Alltagskontexten erkennen.
st22	Die Gültigkeit eines mathematischen Modells für eine reale Situation diskutieren.
st28	Ergebnisse eines mathematischen Modells an der Ausgangssituation prüfen.
st33	Reale Probleme mit mathematischen Mitteln bearbeiten.
<i>Skala: mathematische Darstellungen verwenden</i>	
st04	Tabellen oder Diagramme interpretieren.
st10	Zusammenhänge in Graphen darstellen.
st16	Beziehungen zwischen Tabellen, Graphen oder Diagrammen erkennen.
st23	Für eine Aufgabe die geeignete Darstellungsform (z.B. Tabellen, Gleichungen) auswählen.

st29 Eine geeignete Tabelle oder ein Diagramm selbständig erstellen.

Skala: technisch Arbeiten

st05 Mit Variablen, Gleichungen oder Funktionen arbeiten.
 st11 Standardverfahren wiederholt üben.
 st17 Den Taschenrechner sinnvoll als Hilfsmittel verwenden.
 st24 Mathematische Sätze, Regeln oder Formeln anwenden.
 st30 Mathematische Aufgaben mit dem Computer lösen.

Skala: mathematisch Kommunizieren

st06 Ergebnisse verständlich schriftlich darstellen.
 st12 Eigene Lösungswege anderen Schülern erläutern.
 st18 Erläuterungen in adressatengerechter Fachsprache formulieren.
 st25 Mathematische Inhalte aus Texten selbst erarbeiten.
 st35 Sich mathematische Inhalte gegenseitig erklären.

Anmerkung: Einleitungssatz Lehrkräfte: „Wie oft sollten die Schülerinnen und Schüler in den vergangenen 6 Monaten in Ihrem Unterricht....“; Einleitungssatz Schülerinnen und Schüler: Frage „Wie oft habt ihr in deiner Klasse in den vergangenen 6 Monaten im Mathematikunterricht Folgendes ausgeführt?“

Items im Fragebogen zum mathematischen Selbstkonzept (Schülerinnen und Schüler)

Itemname	Itemformulierung
mathe01_r	Ich bin einfach nicht gut in Mathematik.
mathe02	Im Fach Mathematik bekomme ich gute Noten.
mathe03	In Mathematik lerne ich schnell.
mathe04	Ich war schon immer überzeugt, dass Mathematik eines meiner besten Fächer ist.
mathe05	Im Mathematikunterricht verstehe ich sogar die schwierigsten Aufgaben.

Items im Fragebogen zum mathematischen Interesse (nur für Schülerinnen und Schüler)

Itemname	Itemformulierung
mathe06	Wenn ich mich mit Mathematik beschäftige, vergesse ich manchmal alles um mich herum.
mathe07	Mathematik ist mir persönlich wichtig.
mathe08	Weil mir die Beschäftigung mit Mathematik Spaß macht, würde ich das nicht gerne aufgeben.
mathe09	Für Mathematik interessiere ich mich.
mathe10	Die Beschäftigung mit Mathematik gehört zu meinen Lieblingstätigkeiten.

Items im Fragebogen „Lehrerwahrnehmung (Lehrerbewertung)“ (Schülerinnen und Schüler)

Itemname	Itemformulierung (Antworte so, wie du deinen Mathematiklehrer/ deine Mathematiklehrerin erlebst.)
lehrer01_r	Ich hätte gern einen anderen Mathematiklehrer/ eine andere Mathematiklehrerin.
lehrer02	Mit unserem Mathematiklehrer/ unsere Mathematiklehrerin bin ich sehr zufrieden.
lehrer03	Ich mag unseren Mathematiklehrer/ unsere Mathematiklehrerin sehr gern.
lehrer04	Unser Mathematiklehrer/ unsere Mathematiklehrerin ist selbst vom Fach Mathematik begeistert.
lehrer05	Unser Mathematiklehrer/ unsere Mathematiklehrerin unterrichtet mit Begeisterung.
lehrer06	Unserem Mathematiklehrer/ unserer Mathematiklehrerin scheint das Unterrichten großen Spaß zu machen.

Anhang B: Messabstände in Monaten der Studie 1

IDSchule	zeit1	zeit2	zeit3
1	0	9	14
2	1	8	13
3	0	9	14
4	0	9	14
5	1	8	13
6	0	9	14
7	0	10	14
8	0	9	14
9	0	10	14
10	1	8	13
11	0	10	14
12	1	8	13
13	0	9	14
14	1	8	13
15	1	8	13
16	1	8	13
17	1	9	13

Anmerkung. Den Referenzzeitpunkt bildet der Monat Februar 2007.

Messabstände in Monaten der Studie 2

IDSchule	Zeit1	Zeit2	IDSchule	Zeit1	Zeit2
1	4	10	18	1	10
2	4	10	19	3	10
3	4	10	20	4	10
4	4	10	21	2	11
5	5	10	22	3	10
6	4	10	23	0	10
7	4	10	24	1	10
8	4	11	25	4	11
9	4	10	26	3	12
10	4	10	27	3	10
11	5	10	28	4	10
12	4	10	29	1	10
13	4	10	30	1	10
14	4	10	31	4	11
15	4	10	32	2	11
16	4	10	33	4	10
17	4	11	34	3	11

Anmerkung. Den Referenzzeitpunkt bildet der Monat Mai 2008.

Anhang C

Mittelwert (*MW*), Standardabweichung (*SD*) für den individuellen Wachstumsparameter aus individuellen Regressionen der Merkmalsentwicklung als eine Funktion der Zeit

Merkmal	N	Konstante (Ausgangswert)	Steigung (Veränderungsrate)	Korrelation
		<i>MW (SD)</i>	<i>MW (SD)</i>	
Zusammenarbeit in der FK	44	2.41 (0.42)	0.01 (0.03)	- .43**
Kooperation in der FK	44	3.13 (0.57)	0.00 (0.04)	- .49**
math. Argumentieren	43	3.95 (0.91)	0.00 (0.05)	- .66**
math. Probleme lösen	43	4.26 (0.68)	0.00 (0.06)	- .41**
math. Modellieren	43	3.76 (0.94)	0.02 (0.06)	- .56**
Darstellungen verwenden	43	3.86 (0.79)	0.01 (0.05)	- .57**
technisch arbeiten	43	4.71 (0.67)	0.00 (0.05)	- .63**
math. Kommunizieren	43	4.59 (0.94)	0.01 (0.06)	- .77**

Anhang

Anhang D. Nullmodelle Studie 1, Lehrkräfte

			Einstellungen									
		Parameter	positive	negative	Fach- konferenz	Kooperation	Argu- mentieren	Probleme Lösen	Modellieren	Darstellungen Verwenden	technisch Arbeiten	Kommuni- zieren
Feste Effekte												
Ausgangswert	Konstante	γ_{00}	3.535*** (.061)	2.476*** (.052)	2.514*** (.046)	3.071*** (.056)	3.977*** (.074)	4.265*** (.067)	4.024*** (.078)	3.980*** (.068)	4.668*** (.054)	4.678*** (.069)
Zufallseffekte												
	innerhalb Personen	σ^2_{ε}	.230	.129	.106	.086	.240	.263	.316	.253	.152	.266
Level 1	Ausgangswert	σ^2_0	.192***	.145***	.117***	.192***	.300***	.229***	.328***	.248***	.160***	.247***

Anmerkungen. imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

Studie 2, Lehrkräfte

			Einstellungen									
		Parameter	positive	negative	Fach- konferenz	Kooperation	Argu- mentieren	Probleme Lösen	Modellieren	Darstellungen Verwenden	technisch Arbeiten	Kommuni- zieren
Feste Effekte												
Ausgangswert	Konstante	γ_{00}	3.177*** (.069)	2.651*** (.072)	2.508*** (.047)	2.963*** (.063)	4.274*** (.072)	4.426*** (.066)	4.171*** (.076)	3.934*** (.072)	4.616*** (.054)	4.798*** (.061)
Zufallseffekte												
	innerhalb Per- sonen	σ^2_{ϵ}	.229	.154	.208	.133	.257	.256	.320	.287	.202	.321
Level 1	Ausgangswert	σ^2_0	.285***	.357***	.077***	.272***	.310***	.241***	.323***	.286***	.140***	.151***

Anmerkungen. imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001

Anhang

Studie 1, Schülerinnen und Schüler

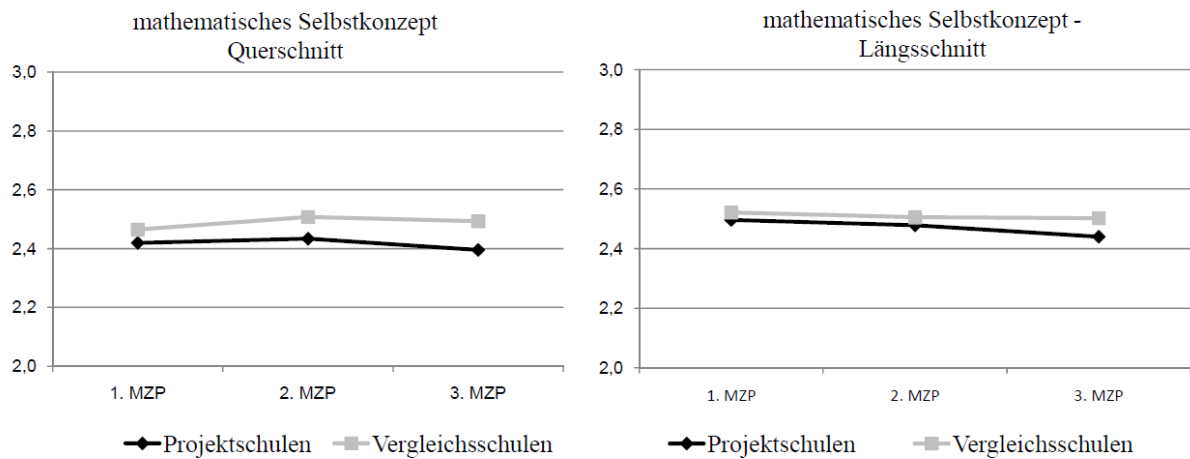
		Parameter	math. Selbst- konzept	Interesse	Lehrerbe- wertung	Argu- mentieren	Probleme Lösen	Modellieren	Darstellungen Verwenden	technisch Arbeiten	Kommuni- zieren	Mathematik- leistung
<i>Feste Effekte</i>												
Ausgangswert	Konstante	γ_{00}	2.465*** (.019)	2.118*** (.017)	2.757*** (.019)	3.822*** (.022)	3.704*** (.020)	3.394*** (.023)	3.401*** (.020)	3.951*** (.017)	3.840*** (.021)	530.417*** (2.703)
<i>Zufallseffekte</i>												
	innerhalb Personen	σ^2_{ε}	.129	.175	.304	.546	.485	.635	.515	.371	.533	2749.716
Level 1	Ausgangswert	σ^2_0	.393***	.297***	.321***	.406***	.291***	.410***	.308***	.209***	.373***	7868.617***
<i>Anmerkungen.</i> imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001												

Studie 2, Schülerinnen und Schüler

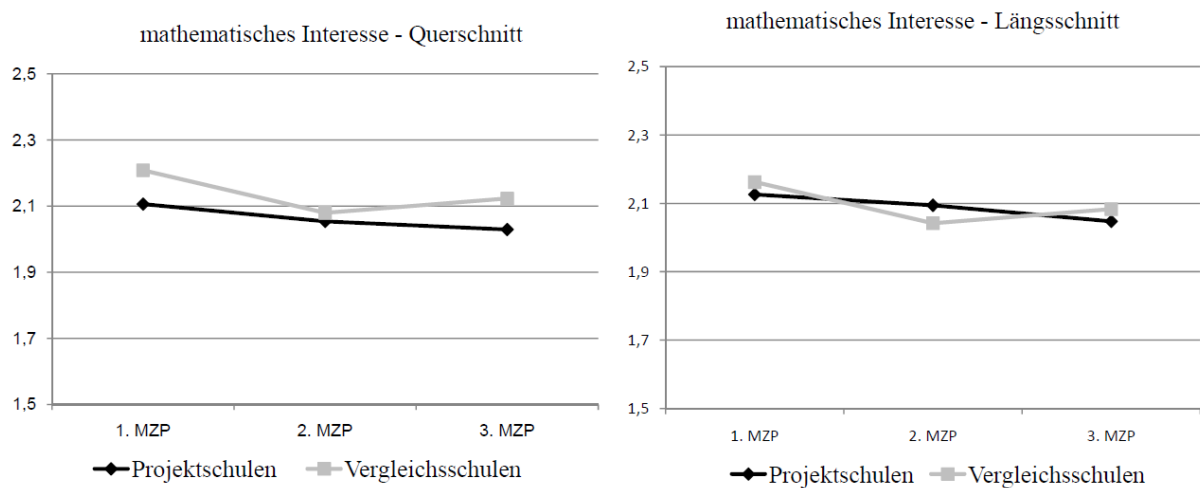
		Parameter	math. Selbst- konzept	Interesse	Lehrerbe- wertung	Argu- mentieren	Probleme Lösen	Modellieren	Darstellungen Verwenden	technisch Arbeiten	Kommuni- zieren	Mathematik- leistung
<i>Feste Effekte</i>												
Ausgangswert	Konstante	γ_{00}	2.493*** (.017)	2.093*** (.015)	2.920*** (.015)	3.926*** (.019)	3.804*** (.018)	3.426*** (.020)	3.405*** (.017)	3.958*** (.015)	3.923*** (.019)	537.578*** (2.319)
<i>Zufallseffekte</i>												
	innerhalb Personen	σ^2_{ε}	.139	.186	.355	.553	.524	.667	.589	.456	.562	2250.526
Level 1	Ausgangswert	σ^2_0	.488***	.327***	.269***	.422***	.350***	.468***	.242***	.202***	.391***	9433.968***
<i>Anmerkungen:</i> imputierter Datensatz mit 5 Imputationen; HLM, Restricted Maximum Likelihood; ~p<.10; *p<.05; **p<.01; ***p<.001												

Anhang E: Abbildungen zu den Entwicklungen aus Schülersicht, Studie 1

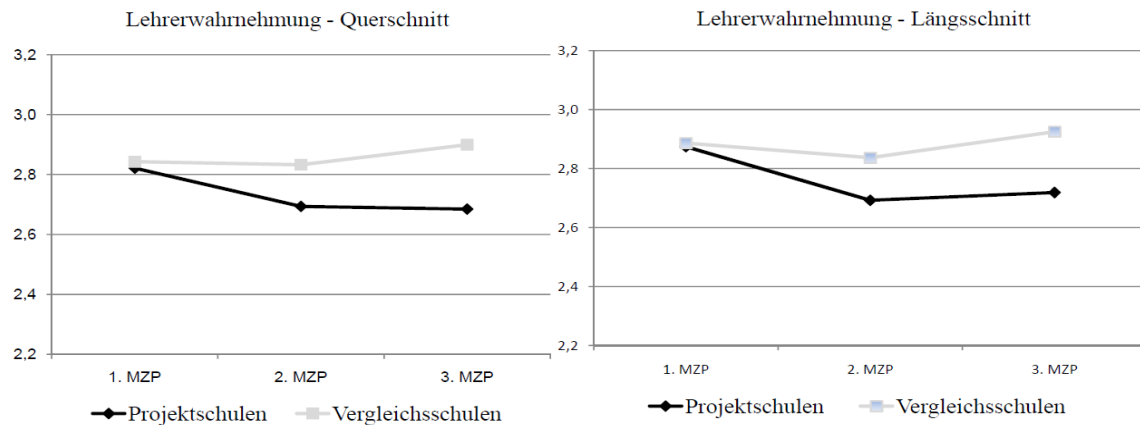
Mathematisches Selbstkonzept im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



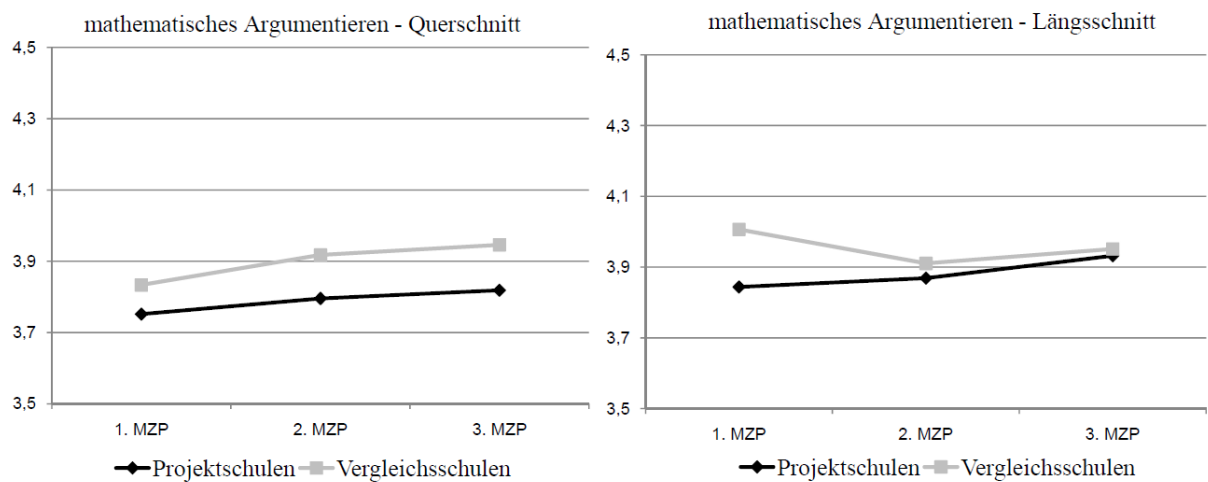
Mathematisches Interesse im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



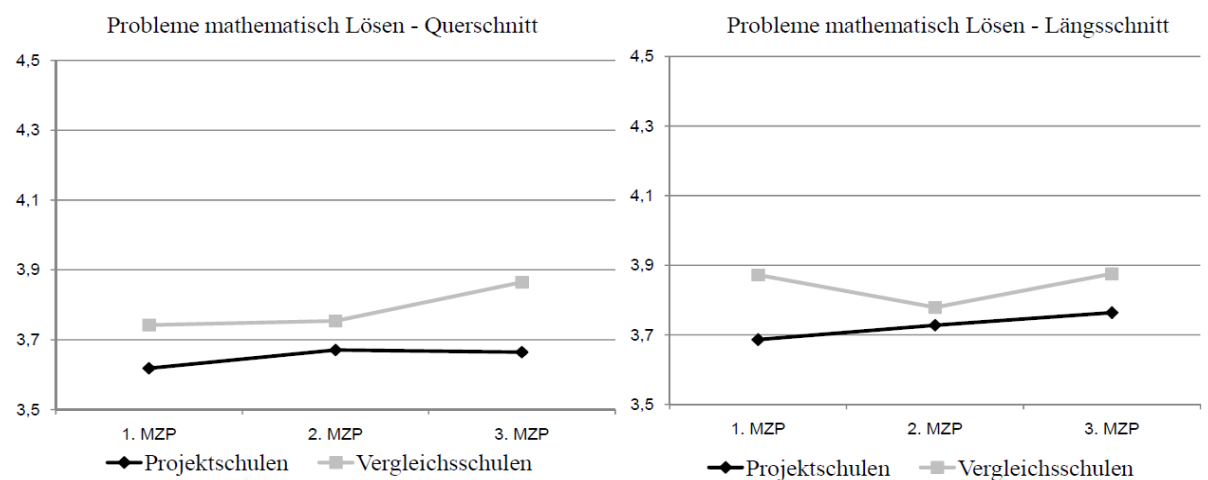
Lehrerwahrnehmung im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



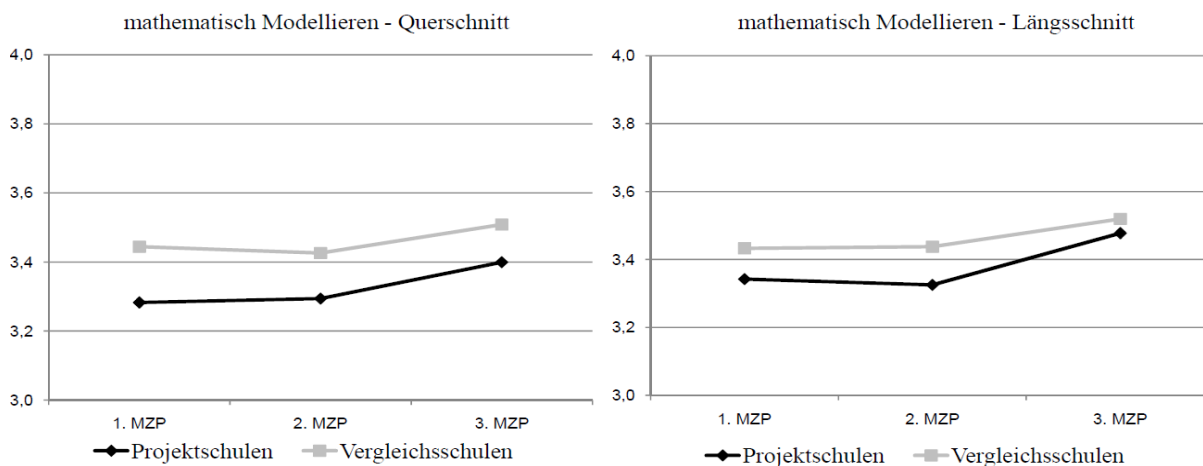
Mathematisches Argumentieren im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



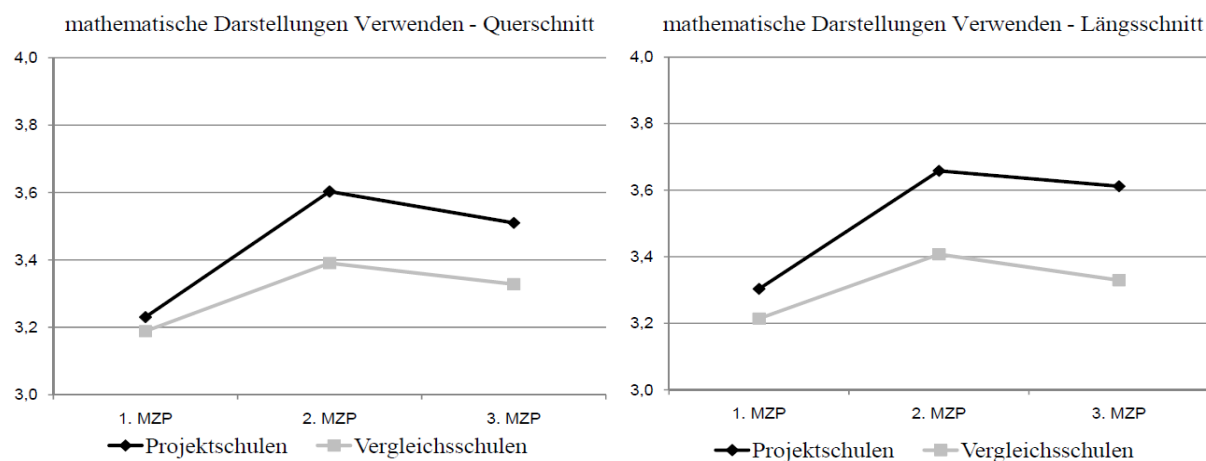
Probleme mathematisch Lösen im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



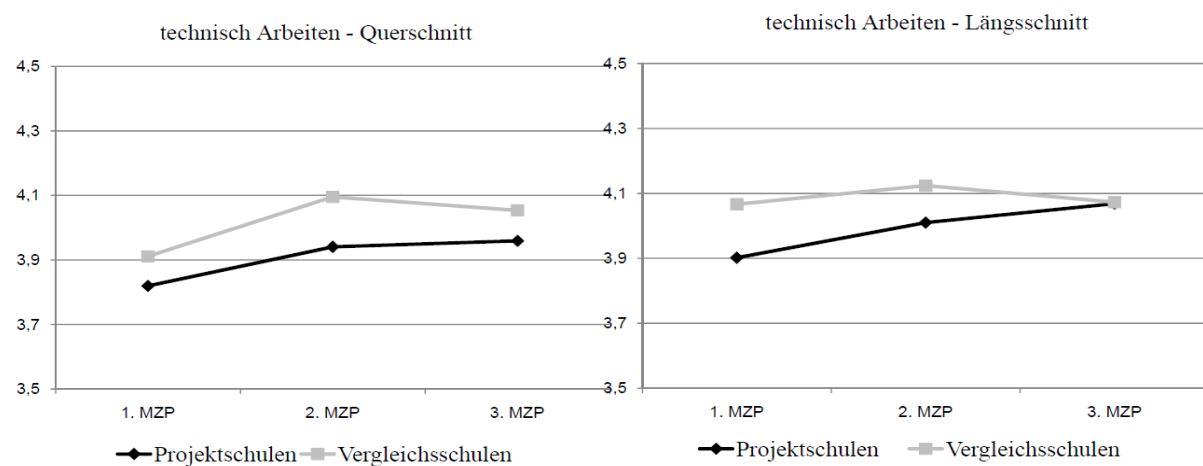
Mathematisch Modellieren im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



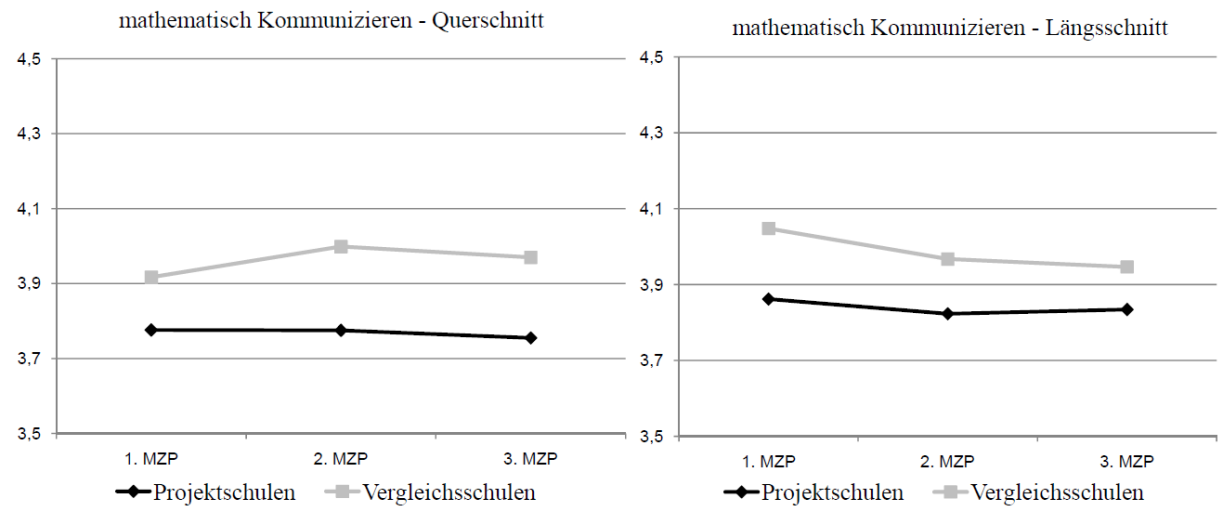
Mathematische Darstellungen verwenden im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



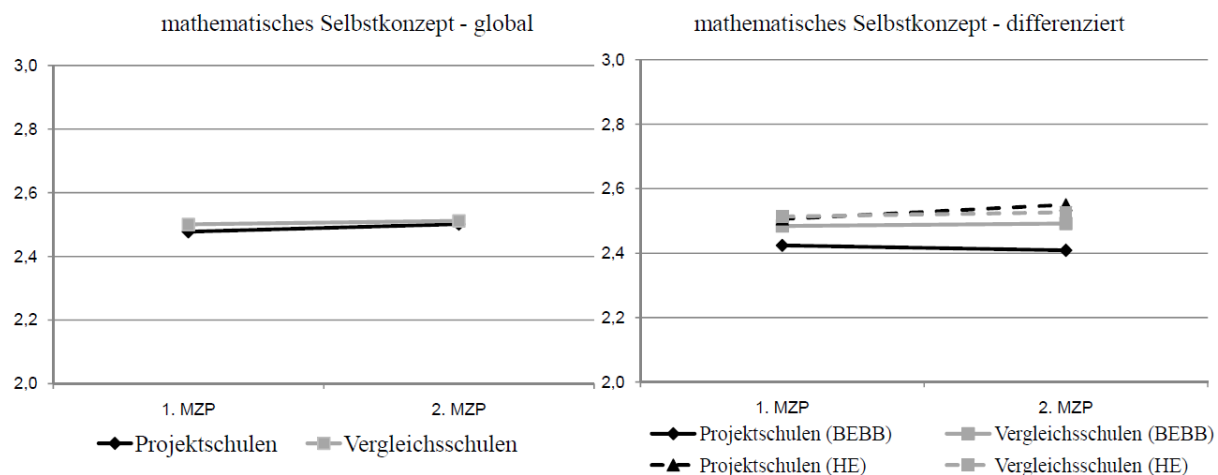
Technisch Arbeiten im Verlauf - Gegenüberstellung aller Teilnehmer und Längsschnitt, Studie 1, Schüler



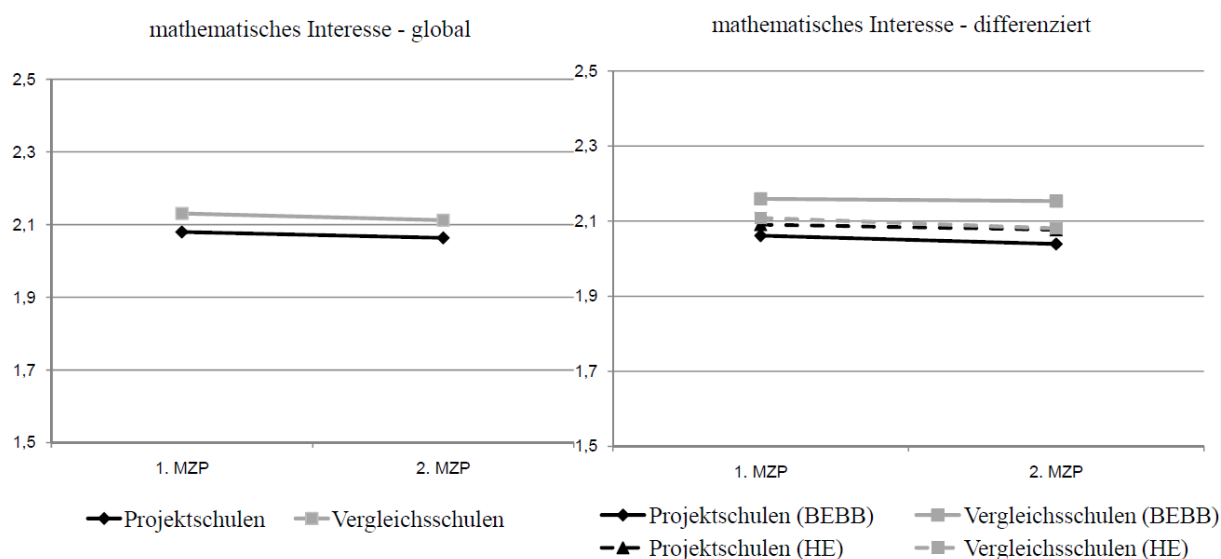
Mathematisch Kommunizieren im Verlauf - Gegenüberstellung aller Teilnehmer und
Längsschnitt, Studie 1, Schüler



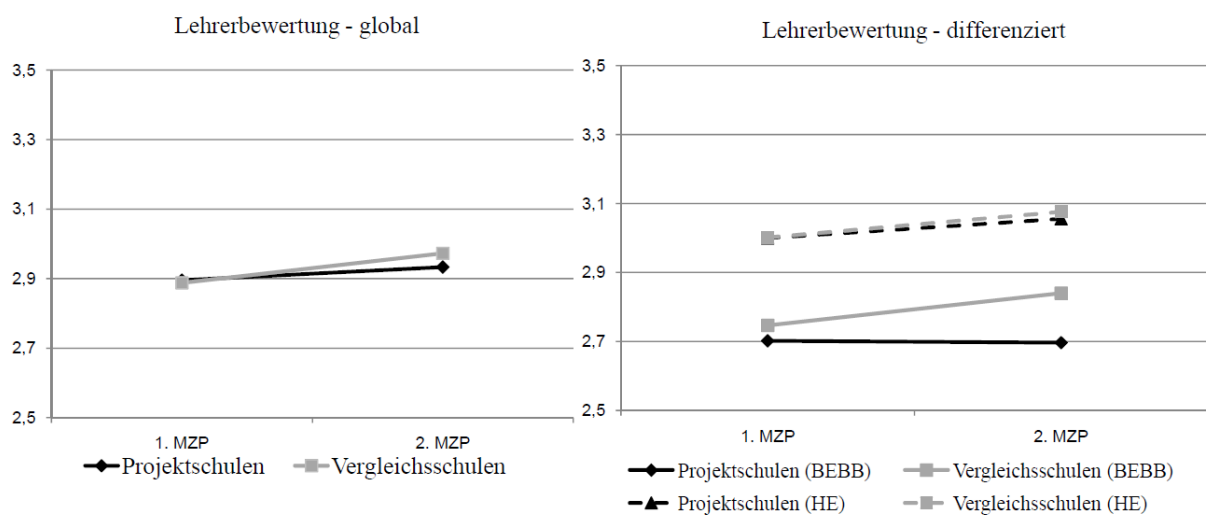
Anhang F: Abbildungen zu den Veränderungen aus Schülersicht, Studie 1
Mathematisches Selbstkonzept in globaler und differenzierter Sicht, Schüler, Studie 2



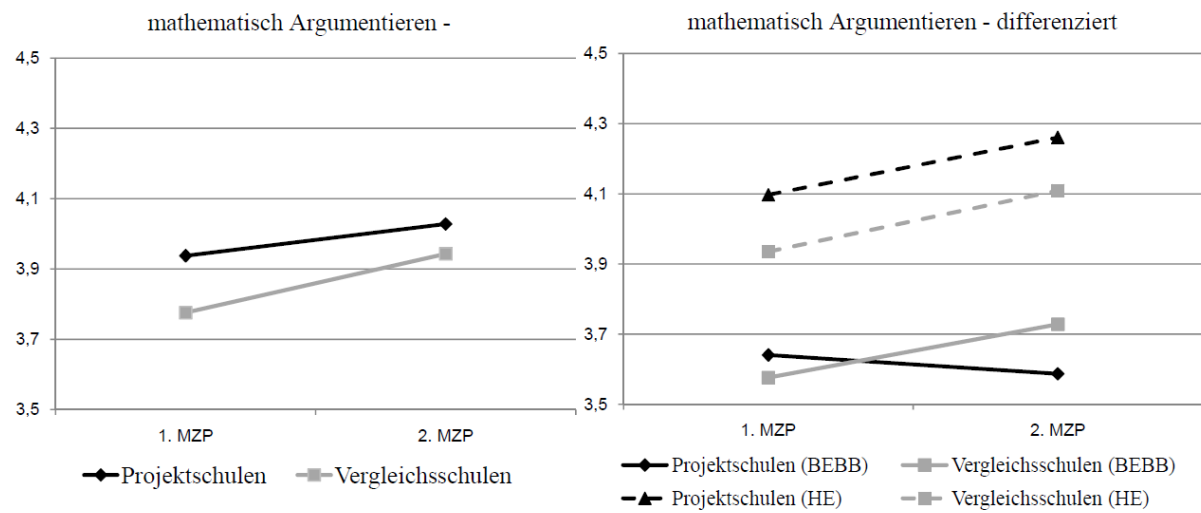
Mathematisches Interesse in globaler und differenzierter Sicht, Schüler, Studie 2



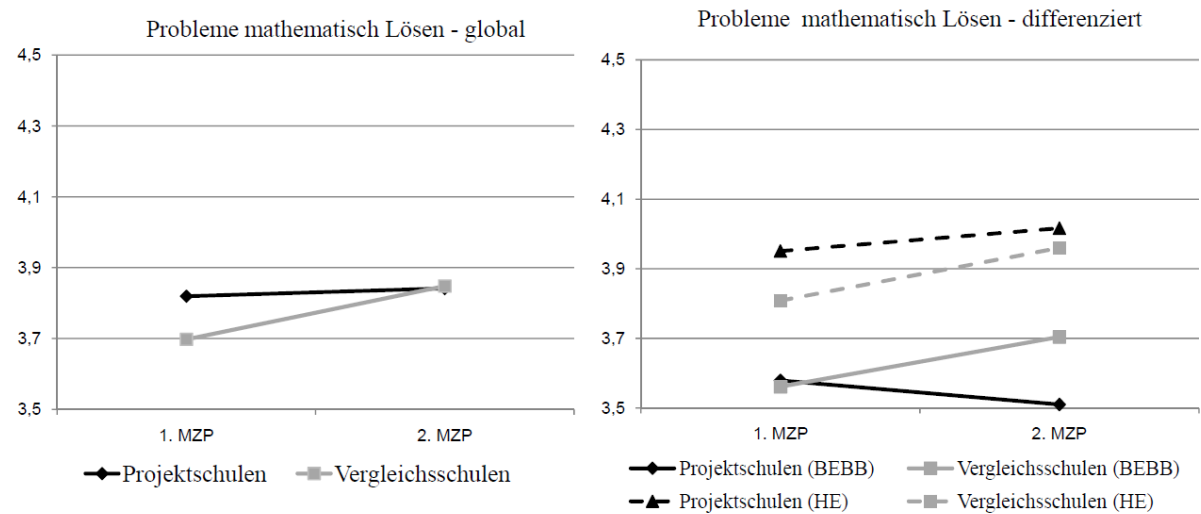
Lehrerwahrnehmung in globaler und differenzierter Sicht, Schüler, Studie 2



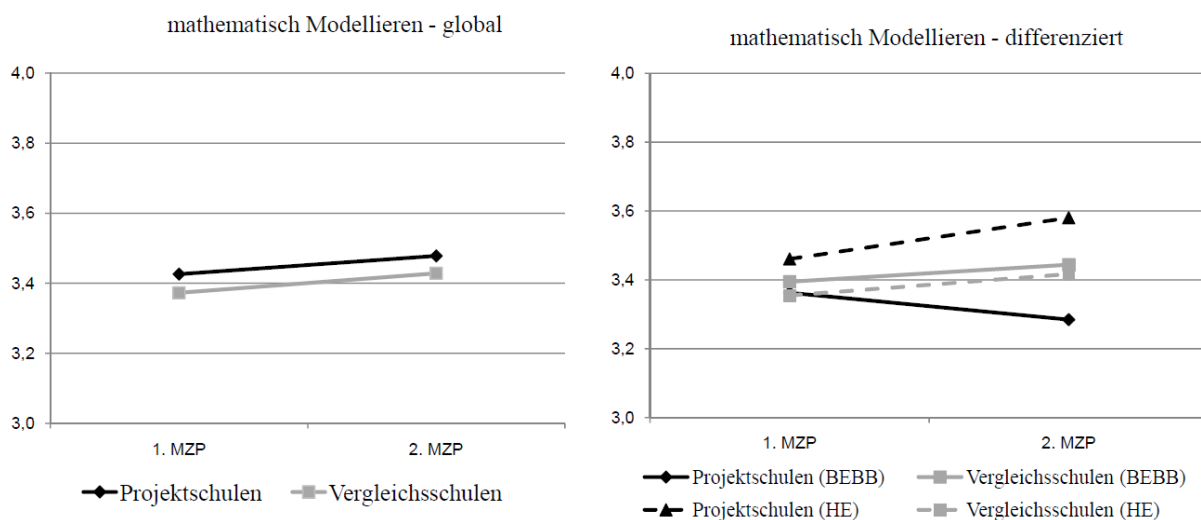
Mathematisch Argumentieren in globaler und differenzierter Sicht, Schüler, Studie 2



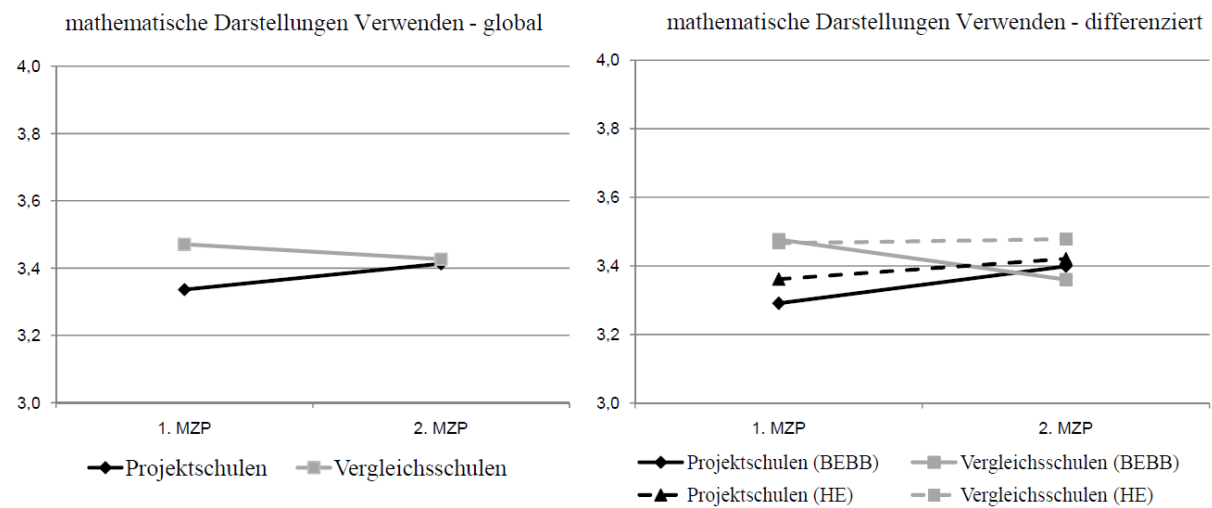
Probleme mathematisch Lösen in globaler und differenzierter Sicht, Schüler, Studie 2



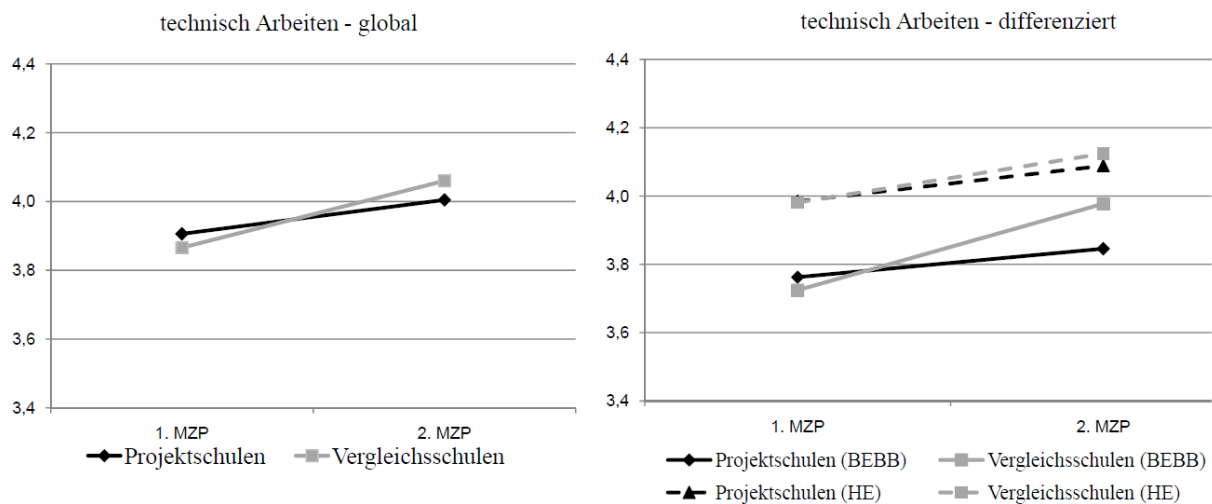
Mathematisch Modellieren in globaler und differenzierter Sicht, Schüler, Studie 2



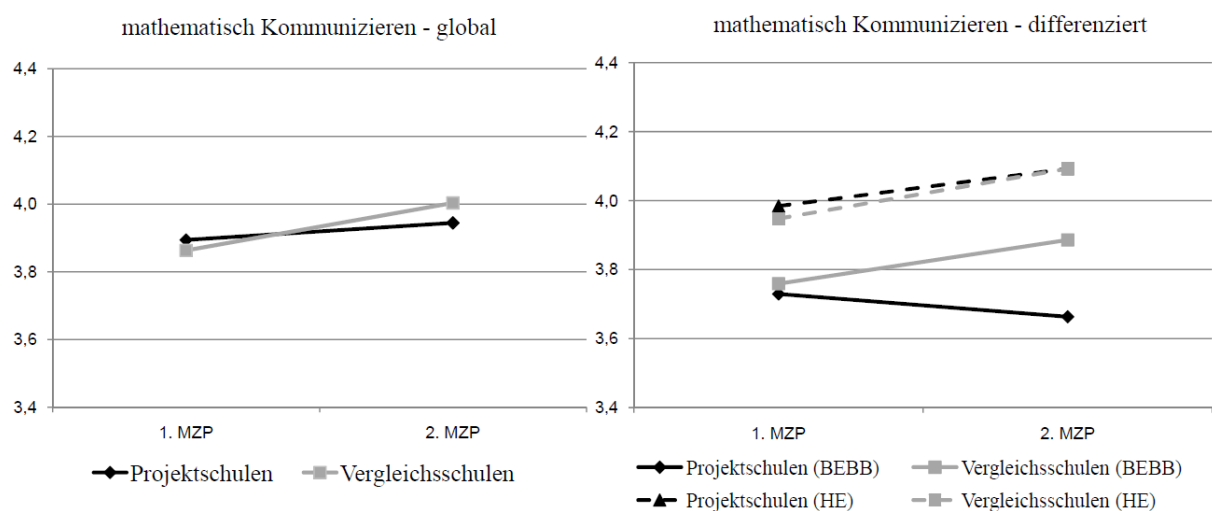
Mathematische Darstellungen Verwenden in globaler und differenzierter Sicht, Schüler, Studie 2



Technisch Arbeiten in Mathematik in globaler und differenzierter Sicht, Schüler, Studie 2



Mathematisch Kommunizieren in globaler und differenzierter Sicht, Schüler, Studie 2



Mathematikleistung in globaler und differenzierter Sicht Schüler, Studie 2

